

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:57:06 ; Search time 13.9441 Seconds
(without alignments)
2740.961 Million cell updates/sec

Title: US-09-602-362E-16

Perfect score: 2599

Sequence: 1 MKVSIPTKALELMDMQTFKA.....IAGDACLQRKNVDVSSTDI 512

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2599	100.0	512	US-09-451-739H-16	Sequence 16, Appl
2	2578	99.2	1002	US-09-620-405B-475	Sequence 475, App
3	2578	99.2	1002	US-09-604-287A-475	Sequence 475, App
4	2578	99.2	1002	US-09-834-759-475	Sequence 475, App
5	2578	99.2	1002	US-09-590-751A-475	Sequence 475, App
6	2578	99.2	1002	US-09-551-621-475	Sequence 475, App
7	2578	99.2	1095	US-09-620-405B-493	Sequence 493, App
8	2578	99.2	1095	US-09-834-759-493	Sequence 493, App
9	2125	81.8	445	US-09-620-405B-473	Sequence 473, App
10	2125	81.8	445	US-09-433-826B-473	Sequence 473, App
11	2125	81.8	445	US-09-604-287A-473	Sequence 473, App
12	2125	81.8	445	US-09-834-759-473	Sequence 473, App
13	2125	81.8	445	US-09-590-751A-473	Sequence 473, App
14	2125	81.8	445	US-09-551-621-473	Sequence 473, App
15	1513.5	58.2	650	US-09-620-405B-469	Sequence 469, App
16	1513.5	58.2	650	US-09-433-826B-469	Sequence 469, App
17	1513.5	58.2	650	US-09-604-287A-469	Sequence 469, App
18	1513.5	58.2	650	US-09-834-759-469	Sequence 469, App
19	1513.5	58.2	650	US-09-590-751A-469	Sequence 469, App
20	1513.5	58.2	650	US-09-551-621-469	Sequence 469, App
21	1513.5	58.2	743	US-09-620-405B-494	Sequence 494, App
22	1513.5	58.2	743	US-09-834-759-494	Sequence 494, App
23	1399	53.8	432	US-09-389-681-181	Sequence 181, App
24	1399	53.8	432	US-09-620-405B-181	Sequence 181, App
25	1399	53.8	432	US-09-339-338-181	Sequence 181, App
26	1399	53.8	432	US-09-433-826B-181	Sequence 181, App
27	1399	53.8	432	US-09-604-287A-181	Sequence 181, App

ALIGNMENTS

RESULT 1
US-09-451-739H-16
; Sequence 16, Application US/09451739H
; Patent No. 6774226
; GENERAL INFORMATION:
; APPLICANT: Jager, Dirk
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gure, Ali
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd
; APPLICANT: Chen, Yao-tseng
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigens
; FILE REFERENCE: LUD 5615
; CURRENT APPLICATION NUMBER: US/09/451,739H
; CURRENT FILING DATE: 1999-11-30
; NUMBER OF SEQ ID NOS: 19
; SEQ ID NO 16
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-451-739H-16

Query Match	100.0.0%;	Score	2599;	DB	4;	Length	512;		
Best Local Similarity	100.0.0%;	Pred. No.	1.7e-179;						
Matches	512;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	MKVSIPTKALELMDMQTFKAEPPEKPSAFEP	1	ESKQDYEESSWDSLSLCTVTSQKDYCVLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM	60				
Db	1	MKVSIPTKALELMDMQTFKAEPPEKPSAFEP	1	ESKQDYEESSWDSLSLCTVTSQKDYCVLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM	60				
Qy	61	ESKQDYEESSWDSLSLCTVTSQKDYCVLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM	120						
Db	61	ESKQDYEESSWDSLSLCTVTSQKDYCVLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM	120						
Qy	121	KVSIPTKALELMDMQTFKAEPPEKPSAFEP	1	ESKQDYEESSWDSLSLCTVTSQKDYCVLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM	180				
Db	121	KVSIPTKALELMDMQTFKAEPPEKPSAFEP	1	ESKQDYEESSWDSLSLCTVTSQKDYCVLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM	180				
Qy	181	SKQKVEENSWDSLSLCTVTSQKDYCVLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM	240						
Db	181	SKQKVEENSWDSLSLCTVTSQKDYCVLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM	240						
Qy	241	RARELQKDHCEQRTGKMEQMKKFCVLKKLSEAKEIKSQLEENQKVKWEQELCSVRLTLN	300						
Db	241	RARELQKDHCEQRTGKMEQMKKFCVLKKLSEAKEIKSQLEENQKVKWEQELCSVRLTLN	300						
Qy	301	QEEKRRNADILNEKIREELGRIEEQHRKLEVEQKQLEEQALRIQDIELKSVESNUNQVSH	360						

```
Db 301 QBEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 360
Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 420
Db 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 420
Qy 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKEILEAEIESHHPRLASAVQDHDQI 480
Db 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKEILEAEIESHHPRLASAVQDHDQI 480
Qy 481 VTSRKSQEPAFHIAAGDACLQKRNKNDVSSTDI 512
Db 481 VTSRKSQEPAFHIAAGDACLQKRNKNDVSSTDI 512

RESULT 2
US-09-620-405B-475
; Sequence 475, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-620-405B-475

Query Match 99.2%; Score 2578; DB 4; Length 1002;
Best Local Similarity 99.6%; Pred. No. 1.3e-177;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADEILPS 60
Db 338 MKVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADEILPS 397
Qy 61 ESKQDYESSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 398 ESKQDYESSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457
Qy 121 KVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADEILPS 180
Db 458 KVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADEILPS 517
Qy 181 SKQKXVEENSWSLSLRETQVSKDVCVPKATHQKEMDKISGKLESDTSLSKILDTVHSC 240
Db 518 SKQKXVEENSWSLSLRETQVSKDVCVPKATHQKEMDKISGKLESDTSLSKILDTVHSC 577
Qy 241 RARELQKDHCCEQRTGMEQMKKFCVLKKLSEAKEIKSQLENOKVWEQELCSVRLTLN 300
Db 578 RARELQKDHCCEQRTGMEQMKKFCVLKKLSEAKEIKSQLENOKVWEQELCSVRLTLN 637
Qy 301 QBEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 360
Db 638 QBEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 697
Qy 361 QBEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 420
Db 697 QBEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 757
Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 420
Db 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 420
```

```
Db 698 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 757
Qy 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKEILEAEIESHHPRLASAVQDHDQI 480
Db 758 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKEILEAEIESHHPRLASAVQDHDQI 817
Qy 481 VTSRKSQEPAFHIAAGDACLQKRNKNDVSST 510
Db 818 VTSRKSQEPAFHIAAGDACLQKRNKNDVSST 847

RESULT 3
US-09-604-287A-475
; Sequence 475, Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-604-287A-475

Query Match 99.2%; Score 2578; DB 4; Length 1002;
Best Local Similarity 99.6%; Pred. No. 1.3e-177;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADEILPS 60
Db 338 MKVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADEILPS 397
Qy 61 ESKQDYESSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 398 ESKQDYESSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457
Qy 121 KVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADEILPS 180
Db 458 KVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADEILPS 517
Qy 181 SKQKXVEENSWSLSLRETQVSKDVCVPKATHQKEMDKISGKLESDTSLSKILDTVHSC 240
Db 518 SKQKXVEENSWSLSLRETQVSKDVCVPKATHQKEMDKISGKLESDTSLSKILDTVHSC 577
Qy 241 RARELQKDHCCEQRTGMEQMKKFCVLKKLSEAKEIKSQLENOKVWEQELCSVRLTLN 300
Db 578 RARELQKDHCCEQRTGMEQMKKFCVLKKLSEAKEIKSQLENOKVWEQELCSVRLTLN 637
Qy 301 QBEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 360
Db 638 QBEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 697
Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 420
Db 698 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 757
Qy 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKEILEAEIESHHPRLASAVQDHDQI 480
Db 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKEILEAEIESHHPRLASAVQDHDQI 480
```

Db 758 KESLTKRASQYSQLKVLIAENTMTLSKLEKQDKKEILEAEIESHHPRLASAVQDHDQI 817

Qy 481 VTSRKSQEPAFHAGDACLQRMNVDVSST 510

Db 818 VTSRKSQEPAFHAGDACLQRMNVDVSST 847

RESULT 4

US-09-834-759-475

; Sequence 475, Application US/09834759

; Patent No. 6680197

; GENERAL INFORMATION:

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Xu, Jiangchun

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.470C9

; CURRENT APPLICATION NUMBER: US/09/834,759

; CURRENT FILING DATE: 2001-04-13

; NUMBER OF SEQ ID NOS: 547

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 475

; LENGTH: 1002

; TYPE: PRT

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (1)...(1002)

; OTHER INFORMATION: Xaa = Any Amino Acid

US-09-834-759-475

Query Match 99.2%; Score 2578; DB 4; Length 1002;

Best Local Similarity 99.6%; Pred. No. 1.3e-177;

Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIEMQKSVPNKALELKNQETLRADIELPS 60

Db 338 MKVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIEMQKSVPNKALELKNQETLRADIELPS 397

Qy 61 ESKQDYEESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120

Db 398 ESKQDYEESSWDSLSLCTVTSQKDVCLPKAXHQKEIDKINGKLEESPDNDGFLKAPCRM 457

Qy 121 KVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIEMQKSVPNKALELKNQETLRADQMPFSE 180

Db 458 KVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIEMQKSVPNKALELKNQETLRADQMPFSE 517

Qy 181 SKQKVVENSWSLSLRETYSQKDVCPKATHQKEMDKISGKLEDSTSLSKILDTVHSCE 240

Db 518 SKQKVVENSWSLSLRETYSQKDVCPKATHQKEMDKISGKLEDSTSLSKILDTVHSCE 577

Qy 241 RARELQDHCBEQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 300

Db 578 RARELQDHCBEQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 637

Qy 301 QEEKRRNADILNEKIREELGRIBEQHRKELEVQQLQEQALRIQDIELKSVESNLNQVSH 360

Db 638 QEEKRRNADILNEKIREELGRIBEQHRKELEVQQLQEQALRIQDIELKSVESNLNQVSH 697

Qy 361 THENENYLLHNCMLKKEIAMLKEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420

Db 698 THENENYLLHNCMLKKEIAMLKEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 757

Qy 421 KESLTKRASQYSQLKVLIAENTMTLSKLEKQDKKEILEAEIESHHPRLASAVQDHDQI 480

Db 758 KESLTKRASQYSQLKVLIAENTMTLSKLEKQDKKEILEAEIESHHPRLASAVQDHDQI 817

Qy 481 VTSRKSQEPAFHAGDACLQRMNVDVSST 510

Db 818 VTSRKSQEPAFHAGDACLQRMNVDVSST 847

RESULT 5

US-09-590-751A-475

; Sequence 475, Application US/09590751A

; Patent No. 6756477

; GENERAL INFORMATION:

; APPLICANT: Yuqiu, Jiang

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Xu, Jiangchun

; APPLICANT: Harlocker, Susan L.

; TITLE OF INVENTION: COMPOSITIONS FOR THE THERAPY AND

; FILE REFERENCE: 210121.470C6

; CURRENT APPLICATION NUMBER: US/09/590,751A

; CURRENT FILING DATE: 2000-06-08

; NUMBER OF SEQ ID NOS: 479

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 475

; LENGTH: 1002

; TYPE: PRT

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (1)...(1002)

; OTHER INFORMATION: Xaa = Any Amino Acid

US-09-590-751A-475

Query Match 99.2%; Score 2578; DB 4; Length 1002;

Best Local Similarity 99.6%; Pred. No. 1.3e-177;

Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIEMQKSVPNKALELKNQETLRADIELPS 60

Db 338 MKVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIEMQKSVPNKALELKNQETLRADIELPS 397

Qy 61 ESKQDYEESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120

Db 398 ESKQDYEESSWDSLSLCTVTSQKDVCLPKAXHQKEIDKINGKLEESPDNDGFLKAPCRM 457

Qy 121 KVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIEMQKSVPNKALELKNQETLRADQMPFSE 180

Db 458 KVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIEMQKSVPNKALELKNQETLRADQMPFSE 517

Qy 181 SKQKVVENSWSLSLRETYSQKDVCPKATHQKEMDKISGKLEDSTSLSKILDTVHSCE 240

Db 518 SKQKVVENSWSLSLRETYSQKDVCPKATHQKEMDKISGKLEDSTSLSKILDTVHSCE 577

Qy 241 RARELQDHCBEQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 300

Db 578 RARELQDHCBEQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 637

Qy 301 QEEKRRNADILNEKIREELGRIBEQHRKELEVQQLQEQALRIQDIELKSVESNLNQVSH 360

Db 638 QEEKRRNADILNEKIREELGRIBEQHRKELEVQQLQEQALRIQDIELKSVESNLNQVSH 697

Qy 361 THENENYLLHNCMLKKEIAMLKEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420

Db 698 THENENYLLHNCMLKKEIAMLKEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 757

Qy 421 KESLTKRASQYSQLKVLIAENTMTLSKLEKQDKKEILEAEIESHHPRLASAVQDHDQI 480

Db 758 KESLTKRASQYSQLKVLIAENTMTLSKLEKQDKKEILEAEIESHHPRLASAVQDHDQI 817

Qy 481 VTSRKSQEPAFHAGDACLQRMNVDVSST 510

Db 818 VTSRKSQEPAFHAGDACLQRMNVDVSST 847

RESULT 6

Tue Feb 22 09:49:10 2005

US-09-551-621-475
; Sequence 475, Application US/09551621
; Patent No. 6825175
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C5
; CURRENT APPLICATION NUMBER: US/09/551.621
; CURRENT FILING DATE: 2000-04-17
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-551-621-475

Query Match 99.2%; Score 2578; DB 4; Length 1002;
Best Local Similarity 99.6%; Pred. No. 1.3e-177; Indels 0; Gaps 0;
Matches 508; Conservative 0; Mismatches 2;

Qy 1 MKVSTPTKALELMDMTFKAEPPKPSAFEPALAEIEMQKSVNKALELKNQOTLRADILPS 60
Db 338 MKVSTPTKALELMDMTFKAEPPKPSAFEPALAEIEMQKSVNKALELKNQOTLRADILPS 397

Qy 61 ESQKDYESSWDSLSCTETVSQKDVCLPKATHQKIDKINGKLESPDNDGFLKAPCRM 120
Db 398 ESQKDYESSWDSLSCTETVSQKDVCLPKATHQKIDKINGKLESPDNDGFLKAPCRM 457

Qy 121 KVSIPTKALELMDMTFKAEPPKPSAFEPALAEIEMQKSVNKALELKNQOTLRADQMPSE 180
Db 458 KVSIPTKALELMDMTFKAEPPKPSAFEPALAEIEMQKSVNKALELKNQOTLRADQMPSE 517

Qy 181 SKQKVEENSWDSLSRETQVSKDVCPKATHQKEMDKISGLKEDSTSLKILDTVHSC 240
Db 518 SKQKVEENSWDSLSRETQVSKDVCPKATHQKEMDKISGLKEDSTSLKILDTVHSC 577

Qy 241 RARELQKHCEQRTGMEQMKKFCVLKKLSEAKIISQLENQKVKWEQELCSVRLTLN 300
Db 578 RARELQKHCEQRTGMEQMKKFCVLKKLSEAKIISQLENQKVKWEQELCSVRLTLN 637

Qy 301 QEEKERNADILNEKIREELGRIEEQRHKELEVKQLEQALRIQDIELKSVESNLNQVSH 360
Db 638 QEEKERNADILNEKIREELGRIEEQRHKELEVKQLEQALRIQDIELKSVESNLNQVSH 697

Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 420
Db 698 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 757

Qy 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEQDKKEILEAEIESHHPRLASAVQDHDQI 480
Db 758 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEQDKKEILEAEIESHHPRLASAVQDHDQI 817

Qy 481 VTSRKSQEPAPHIAGDACLQKKNVDVSST 510
Db 818 VTSRKSQEPAPHIAGDACLQKKNVDVSST 847

RESULT 7
US-09-620-405B-493
; Sequence 493, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu

; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620.405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 493
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(1095)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-620-405B-493

Query Match 99.2%; Score 2578; DB 4; Length 1095;
Best Local Similarity 99.6%; Pred. No. 1.4e-177; Indels 0; Gaps 0;
Matches 508; Conservative 0; Mismatches 2;

Qy 1 MKVSIPTKALELMDMTFKAEPPKPSAFEPALAEIEMQKSVNKALELKNQOTLRADILPS 60
Db 431 MKVSIPTKALELMDMTFKAEPPKPSAFEPALAEIEMQKSVNKALELKNQOTLRADILPS 490

Qy 61 ESQKDYESSWDSLSCTETVSQKDVCLPKATHQKIDKINGKLESPDNDGFLKAPCRM 120
Db 491 ESQKDYESSWDSLSCTETVSQKDVCLPKATHQKIDKINGKLESPDNDGFLKAPCRM 550

Qy 121 KVSIPTKALELMDMTFKAEPPKPSAFEPALAEIEMQKSVNKALELKNQOTLRADQMPSE 180
Db 551 KVSIPTKALELMDMTFKAEPPKPSAFEPALAEIEMQKSVNKALELKNQOTLRADQMPSE 610

Qy 181 SKQKVEENSWDSLSRETQVSKDVCPKATHQKEMDKISGLKEDSTSLKILDTVHSC 240
Db 611 SKQKVEENSWDSLSRETQVSKDVCPKATHQKEMDKISGLKEDSTSLKILDTVHSC 670

Qy 241 RARELQKHCEQRTGMEQMKKFCVLKKLSEAKIISQLENQKVKWEQELCSVRLTLN 300
Db 671 RARELQKHCEQRTGMEQMKKFCVLKKLSEAKIISQLENQKVKWEQELCSVRLTLN 730

Qy 301 QEEKERNADILNEKIREELGRIEEQRHKELEVKQLEQALRIQDIELKSVESNLNQVSH 360
Db 731 QEEKERNADILNEKIREELGRIEEQRHKELEVKQLEQALRIQDIELKSVESNLNQVSH 790

Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 420
Db 791 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 850

Qy 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEQDKKEILEAEIESHHPRLASAVQDHDQI 480
Db 851 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEQDKKEILEAEIESHHPRLASAVQDHDQI 910

Qy 481 VTSRKSQEPAPHIAGDACLQKKNVDVSST 510
Db 911 VTSRKSQEPAPHIAGDACLQKKNVDVSST 940

RESULT 8
US-09-834-759-493
; Sequence 493, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.

APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
FILE REFERENCE: 210121.470C9
CURRENT APPLICATION NUMBER: US/09/834,759
CURRENT FILING DATE: 2001-04-13
NUMBER OF SEQ ID NOS: 547
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 493
LENGTH: 1095
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variant
LOCATION: (1)...(1095)
OTHER INFORMATION: Xaa = Any amino acid
US-09-834-759-493

Query Match 99.2%; Score 2578; DB 4; Length 1095;
Best Local Similarity 99.6%; Pred. No. 1.4e-177;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMTFKABPPKPSAFEPALIEWQKSVPNKALELKNQETLRADIELPS 60
Db 431 MKVSIPTKALELMDMTFKABPPKPSAFEPALIEWQKSVPNKALELKNQETLRADIELPS 490
Qy 61 ESKQDYESSWDSSESLCETVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 491 ESKQDYESSWDSSESLCETVSKQDVCLPKAXHQKEIDKINGKLEESPDNDGFLKAPCRM 550
Qy 121 KVSIPTKALELMDMTFKABPPKPSAFEPALIEWQKSVPNKALELKNQETLRADMPFSE 180
Db 551 KVSIPTKALELMDMTFKABPPKPSAFEPALIEWQKSVPNKALELKNQETLRADMPFSE 610
Qy 181 SKQKVEENSWDSSESLRETYSQKDVCPKATHQKEMDKISGKLESDSTSLKILDTVHSC 240
Db 611 SKQKVEENSWDSSESLRETYSQKDVCPKATHQKEMDKISGKLESDSTSLKILDTVHSC 670
Qy 241 RARELQDHCEQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVWEQELCSVRLTLN 300
Db 671 RARELQDHCEQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVWEQELCSVRLTLN 730
Qy 301 QEEKRRNADILNKEIREELGRIEEQRKELEVKQLEQALRIQDIELKSVESNLNQVSH 360
Db 731 QEEKRRNADILNKEIREELGRIEEQRKELEVKQLEQALRIQDIELKSVESNLNQVSH 790
Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMT 420
Db 791 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMT 850
Qy 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKODKEILAEIESHPRLASAVQDDHDI 480
Db 851 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKODKEILAEIESHPRLASAVQDDHDI 910
Qy 481 VTSRKSQEPAFHIAAGDAQLQRMNVDSST 510
Db 911 VTSRKSQEPAFHIAAGDAQLQRMNVDSST 940

RESULT 9
US-09-620-405B-473
Sequence 473, Application US/09620405B
Patent No. 6528054
GENERAL INFORMATION:
APPLICANT: Jiang, Yuqui
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER

FILE REFERENCE: 210121.470C8
CURRENT APPLICATION NUMBER: US/09/620.405B
CURRENT FILING DATE: 2000-07-20
NUMBER OF SEQ ID NOS: 495
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 473
LENGTH: 445
TYPE: PRT
ORGANISM: Homo sapiens
US-09-620-405B-473
Query Match 81.8%; Score 2125; DB 4; Length 445;
Best Local Similarity 99.8%; Pred. No. 2e-145;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMTFKABPPKPSAFEPALIEWQKSVPNKALELKNQETLRADIELPS 60
Db 26 MKVSIPTKALELMDMTFKABPPKPSAFEPALIEWQKSVPNKALELKNQETLRADIELPS 85
Qy 61 ESKQDYESSWDSSESLCETVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 86 ESKQDYESSWDSSESLCETVSKQDVCLPKAAHQKEIDKINGKLEESPDNDGFLKAPCRM 145
Qy 121 KVSIPTKALELMDMTFKABPPKPSAFEPALIEWQKSVPNKALELKNQETLRADMPFSE 180
Db 146 KVSIPTKALELMDMTFKABPPKPSAFEPALIEWQKSVPNKALELKNQETLRADMPFSE 205
Qy 181 SKQKVEENSWDSSESLRETYSQKDVCPKATHQKEMDKISGKLESDSTSLKILDTVHSC 240
Db 206 SKQKVEENSWDSSESLRETYSQKDVCPKATHQKEMDKISGKLESDSTSLKILDTVHSC 265
Qy 241 RARELQDHCEQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVWEQELCSVRLTLN 300
Db 266 RARELQDHCEQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVWEQELCSVRLTLN 325
Qy 301 QEEKRRNADILNKEIREELGRIEEQRKELEVKQLEQALRIQDIELKSVESNLNQVSH 360
Db 326 QEEKRRNADILNKEIREELGRIEEQRKELEVKQLEQALRIQDIELKSVESNLNQVSH 385
Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMT 417
Db 386 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMT 442

RESULT 10
US-09-433-826B-473
Sequence 473, Application US/09433826B
Patent No. 6579973
GENERAL INFORMATION:
APPLICANT: Jiang, Yuqui
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.470C4
CURRENT APPLICATION NUMBER: US/09/433.826B
CURRENT FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 474
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 473
LENGTH: 445
TYPE: PRT
ORGANISM: Homo sapiens
US-09-433-826B-473

Query Match 81.8%; Score 2125; DB 4; Length 445;
Best Local Similarity 99.8%; Pred. No. 2e-145;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMTFKABPPKPSAFEPALIEWQKSVPNKALELKNQETLRADIELPS 60
Db 26 MKVSIPTKALELMDMTFKABPPKPSAFEPALIEWQKSVPNKALELKNQETLRADIELPS 85

Db 26 MKVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQOTLRADIELPS 85
Qy 61 ESKQDYESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 86 ESKQDYESSWDSLSLCTVTSQKDVCLPKAAHQKEIDKINGKLEESPDNDGFLKAPCRM 145
Qy 121 KVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQOTLRADOMFPSE 180
Db 146 KVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQOTLRADOMFPSE 205
Qy 181 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEMDKISGLKEDSTSLSKILDTVHSCE 240
Db 206 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEMDKISGLKEDSTSLSKILDTVHSCE 265
Qy 241 RARELQKDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQBLCSVRLTLN 300
Db 266 RARELQKDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQBLCSVRLTLN 325
Qy 301 QEEKRRNADILNEKIREELGRIEBQHRKELEVQKQLEALRIQDIELKSYESNLNOVSH 360
Db 326 QEEKRRNADILNEKIREELGRIEBQHRKELEVQKQLEALRIQDIELKSYESNLNOVSH 385
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMT 417
Db 386 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMT 442

RESULT 11
US-09-604-287A-473
; Sequence 473, Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 473
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-604-287A-473

Query Match 81.8%; Score 2125; DB 4; Length 445;
Best Local Similarity 99.8%; Pred. No. 2e-145;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQOTLRADIELPS 60
Db 26 MKVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQOTLRADIELPS 85
Qy 61 ESKQDYESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 86 ESKQDYESSWDSLSLCTVTSQKDVCLPKAAHQKEIDKINGKLEESPDNDGFLKAPCRM 145
Qy 121 KVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQOTLRADOMFPSE 180
Db 146 KVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQOTLRADOMFPSE 205
Qy 181 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEMDKISGLKEDSTSLSKILDTVHSCE 240
Db 206 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEMDKISGLKEDSTSLSKILDTVHSCE 265
Qy 241 RARELQKDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQBLCSVRLTLN 300
Db 266 RARELQKDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQBLCSVRLTLN 325
Qy 301 QEEKRRNADILNEKIREELGRIEBQHRKELEVQKQLEALRIQDIELKSYESNLNOVSH 360
Db 326 QEEKRRNADILNEKIREELGRIEBQHRKELEVQKQLEALRIQDIELKSYESNLNOVSH 385
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMT 417
Db 386 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMT 442

Db 266 RARELQKDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQBLCSVRLTLN 325
Qy 301 QEEKRRNADILNEKIREELGRIEBQHRKELEVQKQLEALRIQDIELKSYESNLNOVSH 360
Db 326 QEEKRRNADILNEKIREELGRIEBQHRKELEVQKQLEALRIQDIELKSYESNLNOVSH 385
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMT 417
Db 386 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMT 442

RESULT 12
US-09-834-759-473
; Sequence 473, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 473
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-473

Query Match 81.8%; Score 2125; DB 4; Length 445;
Best Local Similarity 99.8%; Pred. No. 2e-145;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQOTLRADIELPS 60
Db 26 MKVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQOTLRADIELPS 85
Qy 61 ESKQDYESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 86 ESKQDYESSWDSLSLCTVTSQKDVCLPKAAHQKEIDKINGKLEESPDNDGFLKAPCRM 145
Qy 121 KVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQOTLRADOMFPSE 180
Db 146 KVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQOTLRADOMFPSE 205
Qy 181 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEMDKISGLKEDSTSLSKILDTVHSCE 240
Db 206 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEMDKISGLKEDSTSLSKILDTVHSCE 265
Qy 241 RARELQKDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQBLCSVRLTLN 300
Db 266 RARELQKDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQBLCSVRLTLN 325
Qy 301 QEEKRRNADILNEKIREELGRIEBQHRKELEVQKQLEALRIQDIELKSYESNLNOVSH 360
Db 326 QEEKRRNADILNEKIREELGRIEBQHRKELEVQKQLEALRIQDIELKSYESNLNOVSH 385
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMT 417
Db 386 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMT 442

RESULT 13
US-09-590-751A-473
; Sequence 473, Application US/09590751A
; Patent No. 6756477

```

; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C6
; CURRENT APPLICATION NUMBER: US/09/590,751A
; CURRENT FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 473
; LENGTH: 445
; TYPE: PR
; ORGANISM: Homo sapiens
US-09-590-751A-473

Query Match      81.8%; Score 2125; DB 4; Length 445;
Best Local Similarity 99.8%; Pred. No. 2e-145;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFPAIEIMQKSVPNKALELNEQTLRADEILPS 60
Db 26 MKVSIPTKALELMDMOTFKAEPPKPSAFPAIEIMQKSVPNKALELNEQTLRADEILPS 85
Qy 61 ESKQKDYEESSWDSSESLCETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 86 ESKQKDYEESSWDSSESLCETVSQKDVCLPKAAHQKEIDKINGKLEESPDNDGFLKAPCRM 145
Qy 121 KVSIPTKALELMDMOTFKAEPPKPSAFPAIEIMQKSVPNKALELNEQTLRAEQMFPSE 180
Db 146 KVSIPTKALELMDMOTFKAEPPKPSAFPAIEIMQKSVPNKALELNEQTLRAEQMFPSE 205
Qy 181 SKQKVEENSWDSSESLRETYSQKDVCPKATHQKEMDKISGKLESDSTSLSKILDTVHSC 240
Db 206 SKQKVEENSWDSSESLRETYSQKDVCPKATHQKEMDKISGKLESDSTSLSKILDTVHSC 265
Qy 241 RARELQKHCEORTGKMEQMKKFCVLKKLSEAKEIKSOLENOKVWEQELCSVRITLN 300
Db 266 RARELQKHCEORTGKMEQMKKFCVLKKLSEAKEIKSOLENOKVWEQELCSVRITLN 325
Qy 301 QEEKRRNADILNEKIREELGRIEEQHREKLEVKQLEQALRIQDIELKSVESNLNQVSH 360
Db 326 QEEKRRNADILNEKIREELGRIEEQHREKLEVKQLEQALRIQDIELKSVESNLNQVSH 385
Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKKNAELOMT 417
Db 386 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKKNAELOMT 442

RESULT 15
US-09-620-405B-469
; Sequence 469, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PR
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(650)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-620-405B-469

Query Match      58.2%; Score 1513.5; DB 4; Length 650;
Best Local Similarity 99.0%; Pred. No. 4e-101;
Matches 297; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFPAIEIMQKSVPNKALELNEQTLRADEILPS 60
Db 338 MKVSIPTKALELMDMOTFKAEPPKPSAFPAIEIMQKSVPNKALELNEQTLRADEILPS 397
Qy 61 ESKQKDYEESSWDSSESLCETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120

```

Db 398 ESKQDYESSWDSLSCTVSQKDVCLPKAXHQKEIDKINGKLBESPNDGFLKAPCRM 457
Qy 121 KVSIPTKALELMDMOTFKAEPPKPSAFEPALTEMOKSVPNKALELKNBQTLRADQMPFSE 180
Db 458 KVSIPTKALELMDMOTFKAEPPKPSAFEPALTEMOKSVPNKALELKNBQTLRADQMPFSE 517
Qy 181 SKQKVEENSWDSLSRETYSQKDVCPKATHQKEMDKISGKLEDSTSLSKILDTVHSCE 240
Db 518 SKQKVEENSWDSLSRETYSQKDVCPKATHQKEMDKISGKLEDSTSLSKILDTVHSCE 577
Qy 241 RARELQKHCEORTGMEOMKKKFCVLKKLSEAKEIKSOLENQKVKWEQELCSVR-LTL 299
Db 578 RARELQKHCEORTGMEOMKKKFCVLKKLSEAKEIKSOLENQKVKWEQELCSVRFLT 637

Search completed: February 19, 2005, 02:13:26
Job time : 14.9441 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:55:34 ; Search time 62.9274 Seconds
(without alignments)
4166.462 Million cell updates/sec

Title: US-09-602-362E-16

Perfect score: 2599

Sequence: 1 MKVSIPTKALELMDMQTFKA.....TAGDACLQRKXNVDSSTDI 512

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2589	99.6	1341	2 Q9BXX3	Q9bxx3 homo sapien
2	1542	59.3	1011	2 Q9BXX2	Q9bxx2 homo sapien
3	1094	42.1	424	2 Q9NS19	Q9ns19 homo sapien
4	927.5	35.7	1715	2 Q9UP88	Q9up88 homo sapien
5	927	35.7	1710	2 Q9HIQ1	Q9hiq1 homo sapien
6	871	33.5	1080	2 Q6ZRI4	Q6zri4 homo sapien
7	650.5	25.0	718	2 Q9HCD1	Q9hcd1 homo sapien
8	612.5	23.6	1043	2 Q69ZS2	Q69zs2 mus musculu
9	519	20.0	823	2 Q9H0H6	Q9h0h6 homo sapien
10	475	18.3	453	2 Q81ZM7	Q81zm7 homo sapien
11	431.5	16.6	641	2 Q60311	Q60311 homo sapien
12	313	12.0	2006	2 Q7K5Q6	Q7k5q6 plasmodium
13	313	12.0	2019	2 Q7K5Q5	Q7k5q5 plasmodium
14	313	12.0	2055	2 Q8IHP3	Q8ihp3 plasmodium
15	313	12.0	2055	2 Q8T5C7	Q8t5c7 plasmodium
16	294.5	11.3	520	2 Q80W27	Q80w27 mus musculu
17	294.5	11.3	992	1 AN18_HUMAN	Q8ivf6 homo sapien
18	292.5	11.3	381	2 Q15694	Q15694 homo sapien
19	267.5	10.3	292	2 Q68DM0	Q68dm0 homo sapien
20	265	10.2	2954	2 Q42263	Q42263 xenopus lae
21	264	10.2	450	2 Q6PI49	Q6pi49 homo sapien
22	264	10.2	1738	2 Q76329	Q76329 dictyosteli
23	261.5	10.1	2116	1 MYS2_DICDI	P08799 dictyosteli
24	255	9.8	2057	1 MYSN_DROME	Q99323 drosophila
25	251.5	9.7	1305	2 Q9FJ35	Q9fj35 arabidopsis
26	250	9.6	944	1 NUFI1_YEAST	P32380 saccharomyc
27	247.5	9.5	2230	1 GOA4_HUMAN	Q13439 homo sapien
28	247.5	9.5	2253	2 P70012	P70012 xenopus lae
29	247	9.5	1993	2 Q7PF91	Q7pf91 anopheles g
30	247	9.5	2016	2 Q7PFF90	Q7pf90 anopheles g
31	245.5	9.4	1762	2 Q94DC2	Q94dc2 oryza sativ

ALIGNMENTS

RESULT 1

ID	Q9BXX3	PRELIMINARY;	PRT;	1341 AA.
AC	Q9BXX3			
DT	01-JUN-2001	(TrEMBLrel. 17, Created)		
DT	01-JUN-2001	(TrEMBLrel. 17, Last sequence update)		
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)		
DB	Breast cancer antigen NY-BR-1.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21174979; PubMed=11280766;			
RA	Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karbach J.,			
RA	Jaeger E., Knuth A., Old L.J., Chen Y.T.;			
RT	"Identification of a Tissue-specific Putative Transcription Factor in			
RT	Breast Tissue by Serological Screening of a Breast Cancer Library."			
RL	Cancer Res. 61:2055-2061(2001).			
DR	EMBL: AF269087; AAK27325.1;			
DR	GO: GO:0005634; C:nucleus; NAS.			
DR	GO: GO:0005515; F:protein binding; NAS.			
DR	GO: GO:0003700; F:transcription factor activity; NAS.			
DR	GO: GO:0006355; P:regulation of transcription, DNA-dependent; NAS.			
DR	InterPro: IPR002110; ANK.			
DR	InterPro: IPR001969; Pept_Asp_AS.			
DR	Pfam: PF000023; Ank; 6.			
DR	PRINTS: PR01415; ANKYRIN.			
DR	SMART: SM00248; ANK; 6.			
DR	PROSITE: PS50088; ANK_REPEAT; 4.			
DR	PROSITE: PS50297; ANK_REPEAT_REGION; 1.			
DR	PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_1.			
KW	ANK repeat.			
SQ	SEQUENCE 1341 AA; 152776 MW; 33553DDE6FD3A58B CRC64;			
Query Match 99.6%; Score 2589; DB 2; Length 1341;				
Best Local Similarity 100.0%; Pred. No. 2.2e-96;				
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1 MKVSIPTKALELMDMQTFKARPPEKPSAFEPAPAEIMQKSVPNKALELKNQETLRADIELUPS 60			
Db	677 MKVSIPTKALELMDMQTFKARPPEKPSAFEPAPAEIMQKSVPNKALELKNQETLRADIELUPS 736			
Qy	61 ESKQKDYESSWDSSESICETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120			
Db	737 ESKQKDYESSWDSSESICETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 796			
Qy	121 KVSITPTKALELMDMQTFKARPPEKPSAFEPAPAEIMQKSVPNKALELKNQETLRADQMPFSE 180			
Db	797 KVSITPTKALELMDMQTFKARPPEKPSAFEPAPAEIMQKSVPNKALELKNQETLRADQMPFSE 856			
Qy	181 SKQKVKVENSWSLSRETYSQKDVCPKATHQKEMDKISKLEBDSLSKILDTVHSCB 240			

```

Db      857 SKQKVEENSWDSLSRETYSQDVCVPKATHQKEMDKISGLKEDSTLSKILDTVHSCE 916
Qy      241 PARELOKDHCBORTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVQWEQELCSVRLTLN 300
Db      917 RARELOKDHCBORTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVQWEQELCSVRLTLN 976
Qy      301 QEEKRRNADILNEKIRELGRIEHQHKELEVKQLEQALRIODIELKSVESNLNOVSH 360
Db      977 QEEKRRNADILNEKIRELGRIEHQHKELEVKQLEQALRIODIELKSVESNLNOVSH 1036
Qy      361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db      1037 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 1096
Qy      421 KEESLTTRASQVSGQKVLIAENTMLTSLKKEKQDKEILEAEIESHHPRILASAVQDHDQI 480
Db      1097 KEESLTTRASQVSGQKVLIAENTMLTSLKKEKQDKEILEAEIESHHPRILASAVQDHDQI 1156
Qy      481 VTSRKSQBPAPHIAGDACLQKQNVDSST 510
Db      1157 VTSRKSQBPAPHIAGDACLQKQNVDSST 1186

RESULT 2
Q9BX2 ID Q9BX2 PRELIMINARY; PRT; 1011 AA.
AC Q9BX2; MEDLINE=21174979; PubMed=11280766;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE Breast cancer antigen NY-BR-1.1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21174979; PubMed=11280766;
RA Jaeger D., Stockert B., Guere A.O., Scanlan M.J., Karbach J.,
RA Jaeger E., Knuth A., Old L.J., Chen Y.T.;
RT "Identification of a Tissue-specific Putative Transcription Factor in
RT Breast Tissue by Serological Screening of a Breast Cancer Library.";
RL Cancer Res. 61:2055-2061(2001).
DR EMBL; AF269088; AK27326.1; -.
DR HSP; O75832; UOH.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ANK; 3.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS50088; ANK_REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW ANK repeat.
FT NON_TER 1011 1011
SQ SEQUENCE 1011 AA; 114250 MW; 16627D7B218DE438 CRC64;

Query Match 59.3%; Score 1542; DB 2; Length 1011;
Best Local Similarity 50.5%; Pred. No. 2e-54;
Matches 336; Conservative 42; Mismatches 68; Indels 220; Gaps 5;

Qy      2 KVSIPTKALE-LMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQOTLRADIELPS 60
Db      351 QVCIPESMTQKWEINREVEELPEKPSAFKPAVEMQKTVNKAFLKNEQTLRAAQMPFS 410
Qy      61 ESKQKDYESSWDSLSRETYSQDVCVPKATHQKEMDKISGLKEDSTLSKILDTVHSCE 120
Db      411 ESKQKDYESSWDSLSRETYSQDVCVPKATHQKEMDKISGLKEDSTLSKILDTVHSCE 470
Qy      121 KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQOTLRADIELPS 156
Db      471 KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQOTLRADIELPS 530
Qy      157 -----SVNPKALELKNQOTLRADIELPSKQKVEENSWDSLSRETYSQDVC 206
Db      531 LKPKTCGRKVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQOTLRADIELPS 590

```

```

Qy      207 VPKATHQKEMDKISGLKEDSTLSKILDTVHSCE 224
Db      591 LKPKTCGRKVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQOTLRADIELPS 650
Qy      225 -----SVNPKALELKNQOTLRADIELPSKQKVEENSWDSLSRETYSQDVC 224
Db      651 TFSILFGKPTTENSQSTKVBEEDFNLTITKEGATKVTGQQRDIDIGIIRAPQDQTNKMPTE 710
Qy      225 -----SVNPKALELKNQOTLRADIELPSKQKVEENSWDSLSRETYSQDVC 224
Db      711 LGRKEDTKSTDSSEIISVSDTQNYECLPEATYQKEIKTTNGKJEESEKPSHFEPEATEMQ 770
Qy      225 -----DSTLSKILDTVHSCEARELOKDHCBORTGKMEQMKKFCV 266
Db      771 NSVFNKGLEWKNQOTLRADSTLTKILDALPSCGERELKDCNEQITANWEQMKKFCV 830
Qy      267 LKKLSEAKEIKSQLENQKVQWEQELCSVRLTLNQEEKERNADILNEKIRELGRIEHQ 326
Db      831 LKELSEAKEIKSQLENQKVQWEQELCSVRLTLNQEEKERNADILNEKIRELGRIEHQ 885
Qy      327 HRKELEVKQLEQALRIODIELKSVESNLNOVSHTHENENYLLHNCMLKKEIAMLKLEI 386
Db      886 LRKLEVKQLEQALRIODIELKSVESNLNOVSHTHENENYLLHNCMLKKEIAMLKLEI 945
Qy      387 ATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTTRASQVSGQKVLIAENTML 446
Db      946 ATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTTRASQVSGQKVLIAENTML 1005
Qy      447 TSKLKE 452
Db      1006 TSKLKE 1011

```

RESULT 3

```

Q9NS19 ID Q9NS19 PRELIMINARY; PRT; 424 AA.
AC Q9NS19;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE PRED4 protein (Fragment).
GN Name=PRED4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.S.,
RA Toyoda A., Ishii K., Totoki Y., Choi D.K., Soeda E., Ohki M.,
RA Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordieck G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,
RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspo M.L.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL163203; CAB90394.1; -.
FT NON_TER 1 1
FT NON_TER 424 424
SQ SEQUENCE 424 AA; 50038 MW; C6E704795C534625 CRC64;

Query Match 42.1%; Score 1094; DB 2; Length 424;
Best Local Similarity 79.0%; Pred. No. 8.3e-37;
Matches 226; Conservative 19; Mismatches 27; Indels 14; Gaps 2;

Qy      225 DSTLSKILDTVHSCEARELOKDHCBORTGKMEQMKKFCV 284

```

```

Db 2 DSTLSKILDAVPSRCGKELKDHCEQITAKMEQTKNFKCVLQKELSEAKIKSQLENO 61
QY 285 KVKWEQELSVRLTLNQBEKERNADILNEKIREBLGRIEQHRKELEVKQLEQALRIQ 344
Db 62 KAKWEQELCS-----KKRDVDILKEIRP-----EEQLRKKELEVKQLEQALRIQ 107
QY 345 DIELKSVSNNQVSHTHENENYLLHENCMLKKEIAMLKLEIATILKHQYQEKENKYFEDI 404
Db 108 DIELKSVSNNQVSHTHENSENDLPHENCMLKKEIAMLKLEIATILKHQYQEKENKYFEDI 167
QY 405 KILKEKNAELQWTLKKEESLTKRASQYSGOLKVLIAENTMLTSKLKEKQDKEIIEAEIE 464
Db 168 KILQEKNAELQWTLKQKTLTKRASQYREQVKVIAENTMLTSKLKEKQDKEIIEAEIE 227
QY 465 SHHPLASAVQDHDQVITSRKQSEPAFHAGDACIQRKMNVDVST 510
Db 228 SHHPLASALQDHDQVITSRKQELAFHAGDAHLQGMVDVST 273

RESULT 4
QUPUS8 PRELIMINARY; PRT; 1715 AA.
AC QUPUS8;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAR-2004 (TremBLrel. 26, Last sequence update)
DE KIAA1074 protein (Fragment).
GN Name=KIAA1074;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=9397452; PubMed=10470851;
RA Kikuno R., Nagase T., Ishikawa K., Hirose M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIV. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.";
RL DNA Res. 6:197-205(1999).
DR EMBL; AB028997; BAA83026.2; -.
DR HSSP; P20749; 1K1B.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 4.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW ANK repeat.
FT NON TER
SQ SEQUENCE 1 1715 AA; 196962 MW; A82858F5F58E3203 CRC64;

Query Match 35.7%; Score 927.5; DB 2; Length 1715;
Best Local Similarity 40.7%; Pred. No. 1.6e-29;
Matches 233; Conservative 94; Mismatches 156; Indels 89; Gaps 13;

QY 21 EPPEKPSAFEPALIMQKSVNKAELKNEQTLRADE-ILPSESKQDYESSWDSLSCE 79
Db 494 ESPERYHLKPTIEMKDSVPNKAQGMKDVQTSKAAAEHDLVASEEEQREGSENNQPOVE 553
QY 80 TVSQKDVCLPKATHQKEI-----DKINGKLEESPDNDGFL---KAPCRMKVSIPTK---- 127
Db 554 EERKK-----HRNNEVEVSANIHDGATDDAEDDDDDGLIQRKSGETHQOFPKENKE 608
QY 128 -----ALELMDMQ-TFKAEPPKPSAFEPALIMQKSVNKAELKNEQTL-----RA 173
Db 609 YASGPALQMKVEKSTESKESVNSPVFGKASLLTGGLIQVDDSSLSIDEDEGR 668
QY 174 DMFPSESKQKKEVNSWDS-----ESLRETVSQKDVCPKATHQKEM---DKISGLKEDS 226
Db 669 TKKTSNEXNKNVKNQIQSMDDVDDLTQSSETASE-DCELPHSSYKNFMILLIEQLGMECKDS 727

```

```

QY 227 TSLSKILTIVHSCERARBLQKHDCSQRTGKMEQMKKFCVLKKLSEAKIKSQLENOKV 286
Db 728 VSLKIQDAALSCEKRLLEKKNHCCELLTVKIKMEDKVNVLQRELSSETKEIKSQLEHQKV 787
QY 287 KWEQELSVRLTLNQBEKERNADILNEKIREBLGRIEQHRKELEVKQLEQALRIQDI 346
Db 788 EMERELCSRLFSNLQBEKERNADTLYEKIREQLRKEBEQYRKEVEVQQLSLSQTLDEM 847
QY 347 ELKSVESNLNV-----SH 360
Db 848 ELRTVKSNNLVQVQERNDQAQRLSREQVARNMLQDGLITNHLSSKQKEIEMAKKMNSENH 907
QY 361 THENENYLLHENCMLKKEIAMLKLEIATILKHQYQEKENKYFEDIKILKEKNAELQWTLK 420
Db 908 SHEEKOLSHKNSMLQEBIAMLRLEIDITIKNQBEKKEKCPEDLKIYKKNEDLQKTIKQ 967
QY 421 KEESLTKRASQYSGOLKVLIAENTMLTSKLK-EKQDKEIIEAEIESHPRLASAVQDHDQ 479
Db 968 NEETLTQTISQYNGRLSVLTAEANMLSKLENKOSKERLEAEVESHSLAAATHDRDQ 1027
QY 480 IVTSRKQSEPAFHAGDAC--LQRKMNVDVSS 509
Db 1028 SETSKRELELAFQARDECRLQDKMNFVSN 1059

RESULT 5
Q9H101 PRELIMINARY; PRT; 1710 AA.
AC Q9H101;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DE BAI45E8.1 (KJAA1074).
GN Name=BAI45E8.1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bird C.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL162272; CAC19649.1; -.
DR HSSP; P20749; 1K1B.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 4.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW ANK repeat.
SQ SEQUENCE 1710 AA; 196409 MW; 01CBF9BADB894872 CRC64;

Query Match 35.7%; Score 927; DB 2; Length 1710;
Best Local Similarity 40.7%; Pred. No. 1.7e-29;
Matches 233; Conservative 94; Mismatches 156; Indels 90; Gaps 13;

QY 21 EPPEKPSAFEPALIMQKSVNKAELKNEQTLRADE-ILPSESKQDYESSWDSLSCE 79
Db 488 ESPERYHLKPTIEMKDSVPNKAQGMKDVQTSKAAAEHDLVASEEEQREGSENNQPOVE 547
QY 80 TVSQKDVCLPKATHQKEI-----DKINGKLEESPDNDGFL---KAPCRMKVSIPTK---- 127
Db 548 EERKK-----HRNNEVEVSANIHDGATDDAEDDDDDGLIQRKSGETHQOFPKENKE 602
QY 128 -----ALELMDMQ-TFKAEPPKPSAFEPALIMQKSVNKAELKNEQTL-----R 172
Db 603 YASGPALQMKVEKSTESKESVNSPVFGKASLLTGGLIQVDDSSLSIDEDEGR 662
QY 173 ADQMPFSESKQKKEVNSWDS-----ESLRETVSQKDVCPKATHQKEM---DKISGLKED 225
Db 663 TKKTSNEXNKNVKNQIQSMDDVDDLTQSSETASE-DCELPHSSYKNFMILLIEQLGMECKD 721

```



```

Qy 226 STSLKILDTVHSCERARELOKHCEQRTGMEQMKKFCVLKXKLSAKKIKSOLENOK 285
Db 722 SVSLKIQDAALSCERLLEKXKHCHELTATVKKMEDKVNVLQRELSKTKIQLQEHQK 781
Qy 286 KWVEQELCSVRLTNOBEKERNADILNEKIREELGRIEEQRKELEVKQOLEQALRIQD 345
Db 782 VWEERELCSLRSFSLNOBEKERNADILNEKIREELGRIEEQRKELEVKQOLEQALRIQD 841
Qy 346 IELKSVEENLNOV-----S 359
Db 842 MELRTVKSNLNVQVRNDAQRLSREQNAQRLSDGILTNHLSKQKEIEMAKKMSNS 901
Qy 360 HTHENENYLLHNCMLKKEIAMLEIATLKHQYOEKENYFEDIKILKEKNAELQMTLK 419
Db 902 HSHEEKOLSHKNSMLQEBIAMLEIATLKHQYOEKENYFEDIKILKEKNAELQMTLK 961
Qy 420 LKEESITKRAQSYGOLKVLIAENTMTLSKLK-EKODKEILEAEIESHPRLASAVQDHD 478
Db 962 QNEEILTQISYNGRLSVLTAEANMLNSKLENEKQSKERLEAEVESHSLAAIHDRD 1021
Qy 479 QIVTSRKSOEPAFHAGDAC--LQRKNVDVSS 509
Db 1022 QSETSKRELEAFQARDECSRLQDKMNFVSN 1054

RESULT 6
Q6ZR14
ID Q6ZR14 PRELIMINARY; PRT; 1080 AA.
AC Q6ZR14;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE K1AA1641 protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Trachea;
RA Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,
RA Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A.,
RA Tahil S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuna M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK128577; BAC87508.1; -.
FT NON_TER 1080 1080
SQ SEQUENCE 1080 AA; 126741 MW; 47BBAAA933AE595C CRC64;

Query Match 33.5%; Score 871; DB 2; Length 1080;
Best Local Similarity 38.7%; Pred. No. 1.9e-27;
Matches 225; Conservative 97; Mismatches 153; Indels 106; Gaps 14;

Qy 2 KVSIPKALELMDQTFKAE-----PPEKSAFPALEIMQKSVNKALELKN---EOTLRA 54
Db 64 KVSLEKELOQ--DMQRFKNIGMLKVEFOALEKEKVLQKEVEERKGRHNMESVANI 121
Qy 55 DEILPESKQKQVESSWDSSELSCTVTSQKDVCLPK-----ATHQKEIDKINGK 103
Db 122 HDGATDDAEDDDGLGLOKRSGETDHOQ---FPRKENKEYASSGALQMKVEKSTEXE 178
Qy 104 LEESPNDGFLKAPCRMKVSIPTKALELMDQTFKABPPE---KPSAFPALEIMQKSVN 160
Db 179 KRTSKES---VNSPVFGKASLLTGGLIQVDDDSLSLSEIDDEGRPT-----KRTSN 226
Qy 161 KALELKNQETLRADQMPFSSKQKQVESSWDSSELSRETYSQKDVCPKATHQKEM---D 217
Db 227 EKNKVRNQISQMD-----VDDLQSSSETASE-----DCELPHSSYKNFMILLIE 270

```

```

Qy 218 KISKLEDTSLSKILDTVHSCERARELOKHCEQRTGMEQMKKFCVLKXKLSAKKIKSOLENOK 277
Db 271 QLGWECKSVSLKIQDAALSCERLLEKXKHCHELTATVKKMEDKVNVLQRELSKTKIQLQEHQK 330
Qy 278 KSQLENQKQVWEQELCSVRLTNOBEKERNADILNEKIREELGRIEEQRKELEVKQOLEQALRIQD 337
Db 331 KSQLEHQKQVWEQELCSVRLTNOBEKERNADILNEKIREELGRIEEQRKELEVKQOLEQALRIQD 390
Qy 338 EQALRIQDIELKSVEENLNOV-----S 358
Db 391 ELSQTLMELEIRTVKSNLNVQVRNDAQRLSREQNAQRLSDGILTNHLSKQKEIEMAKKMSNS 450
Qy 359 -----SHTHENENYLLHNCMLKKEIAMLEIATLKHQYOEKENYFEDIKILKEKNAELQMTLK 411
Db 451 KKNSENHSHSEBEKOLSHKNSMLQEBIAMLEIATLKHQYOEKENYFEDIKILKEKNAELQMTLK 510
Qy 412 AELQMTLKLKEESITKRAQSYGOLKVLIAENTMTLSKLK-EKODKEILEAEIESHPRLASAVQDHD 478
Db 511 EDLQKTIKQNEEILTQISYNGRLSVLTAEANMLNSKLENEKQSKERLEAEVESHSLAAIHDRD 570
Qy 471 ASAVODHDQIVTSRKSOEPAFHAGDAC--LQRKNVDVSS 509
Db 571 AAAIHDRQSETSKRELEAFQARDECSRLQDKMNFVSN 611

RESULT 7
Q9HCD1
ID Q9HCD1 PRELIMINARY; PRT; 718 AA.
AC Q9HCD1;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE K1AA1641 protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
EX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVIII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro.";
RL DNA Res. 7:273-281(2000).
DR EMBL; AB046861; BAB13467.1; -.
FT NON_TER 1
SQ SEQUENCE 718 AA; 82972 MW; 2AD05ADE70E3C97F CRC64;

Query Match 25.0%; Score 650.5; DB 2; Length 718;
Best Local Similarity 34.9%; Pred. No. 9.3e-19;
Matches 177; Conservative 102; Mismatches 185; Indels 43; Gaps 12;

Qy 24 EKPSAFPALEIMQKSVNKALELKNQETLRADQMPFSSKQKQVESSWDSSELSCTVTSQ 83
Db 13 QKQPAKATSKQSVSNIPTEIKDQGO-----SGTVSSQK---QPAWKATSV-----K 58
Qy 84 KDVCPLKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQTFKABPPE 143
Db 59 KDSVSNIAETKD-QQIRGTV--SPQKSAQKVPFKKVVLLNLIATRITGGWKSCTEYPE 115
Qy 144 KPSAFPALEIMQKSVNKALELKNQETLRADQMPFSSKQKQVESSWDSSELSCTVTSQ 201
Db 116 NLPTLKATIENTKNSVLNTATKMKDVQSTPTEQDLEMASEGEQKRLSEYENQPOVKYQIH 175
Qy 202 QK-----DVCVPKATHQK-----MDKISGKLEDTSLSKILDTVHSCERAR 243
Db 176 SRDDLDDLIQSSQTSVEDGDSLCCNCKNVILLIDHEMKCKDCVHLKIKNTFCLMKRLI 235
Qy 244 ELQKDHCEQRTGMEQMKKFCVLKXKLSAKKIKSOLENOKVWEQELCSVRLTNOBE 303
Db 236 KLKDNHCEQLRKIRKLNKKNKASVLQKRISKEBEIKSQLKHEILELEKELCSLRAFAQBK 295

```

```

Qy 304 EKRRNADILNEKIRELGRIBEQHRKELEVKQOQLRQALRIODIELKSVESNLNQVSHTHE 363
Db 296 KRRNVEELHOKVREKLRTBEQVRIEADVTPIKPAKLSAEVLKGTGNNNSQVSETDE 355
Qy 364 NENYLLHNCMLKKEIAMKLEIATLKHQYOEKENKYPEDIKILKEKNAELQMTLKLKEE 423
Db 356 KED-LLHNRWQDEIARLEKLEDTIKNQNLK--KYLKDFEIVKRKHEDIQKALKRNGE 412
Qy 424 SLTKRASQYSGQLKVLIAENTMLTSLK--KEKQDKLEIAEIESHPRLASAVQDHDQIVT 482
Db 413 TLAKTIACYSQLAALTIDENTITLSKLEKQESRQRLTEMQSYHCLRNARCDHQSHS 472
Qy 483 SRKQEPAFHAGDAC--LQRKMNVDV 507
Db 473 SKRQELAFQGTVDKCRHLQENLNSHV 499

RESULT 8
Q69ZS2
ID Q69ZS2 PRELIMINARY; PRT; 1043 AA.
AC Q69ZS2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE MKIAA1074 protein (Fragment).
GN Name=MKIAA1074;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Soga Y., Seino S., Nishimura M., Kaisho T., Hoshino K., Kitamura H.,
RA Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene:
RT IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous
RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
RT Randomly Sampled from Size-Fractionated Libraries.";
RL DNA Res. 11:205-218(2004).
DR EMBL; AK173096; BAD32374.1; -.
DR InterPro; IPR010989; t-snare.
DR InterPro; IPR009054; Topismrse_insert.
FT NON_TER 1
FT TER 1043
SQ SEQUENCE 1043 AA; 121354 MW; 5F37B1F9AA37B212 CRC64;

Query Match 23.6%; Score 612.5; DB 2; Length 1043;
Best Local Similarity 32.1%; Pred. No. 4.5e-17;
Matches 177; Conservative 84; Mismatches 143; Indels 147; Gaps 15;

Qy 20 AEPKESAFPEAEMQKSVNKALELKNEQTLR-----ADEILPSESQKQDYESS 71
Db 1 AEP--SGNLYSGAAGDGVAPQSGDTEHQSPREGSGRGGPALLMKEAKGVENEK-- 56
Qy 72 WDSLSCTVTSQKDVCLPKATHQEKIDKINGKLES PDNDGFLKAPCRMKVSIPTKALEL 131
Db 57 WVSREPARTANSERTGLP-----TCGWPMQDGCW----- 87
Qy 132 MDMQTFKAEPEKPSAFPEAEMQKSVNKALELKNEQTLRADQM-----FPSESQKQKVE 187
Db 88 SDTQSEARPTKTS-----SKHNK---DSGQTAAVDNLDLPFESSETASEDHE 133
Qy 188 ENSWDSSELRVTSQKDVCPKATHQEKMDKISGLKEDSTSLSKILDVHSCERARELOK 247
Db 134 LQGPDSESI-----LCA-----IEHLRECKDTASLLKIRDAVYSYKSLIELKR 177
Qy 248 DHCORTGKMBOMKKFCVKKLSEAKEIKSOLENOKVKEQELCSVRLTLNQBEKKRR 307
Db 178 SHCELLTGKLRMENKYGLQKEMSETSEVKSRLSEHEKVGWEQELCULRFAKQEEKRR 237
Qy 308 NADILNEKIRELGRIBEQHRKELEVKQOQLRQALRIODIELKSVESNLNQV----- 358

```

```

Db 238 SADQLSEKTMQLRRKKEQCOQSEVEARQOQLASRLTLEMELKTVKSHLNQVLEERNETQR 297
Qy 359 -----SH-----THENYLLHNCMLKKEIAM 381
Db 298 QLSREQNARMLQDGLASHLCKQKEIEMTQKKMTSEVSVSHEKEKDLHKQRLQDEVAV 357
Qy 382 LKLEIATLKHQYQKBNKYFEDIKILKEKNAELQMTLKEESLTKRASQYSGQLKVLIA 441
Db 358 LRLEMDTIKSHNQKEKRYLEDIKIANEKDNLQRMVKL----- 396
Qy 442 ENTMLTSLK--KEKQDKLEIAEIESHPRLASAVQDHDQIVTSRKSPQAFHAGDAC-- 498
Db 397 --NMLSSKLDNEKQKQRLTVDVSEFRSLASALHDHAEIQTAQRDLLEIAFORARDEWFR 454
Qy 499 LQRKMNVDVSS 509
Db 455 VKDKNFDMSN 465

RESULT 9
Q9H0H6
ID Q9H0H6 PRELIMINARY; PRT; 823 AA.
AC Q9H0H6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp434A171.
GN Name=DKFZp434A171;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TiSSUE=Testis;
RG The German cDNA Consortium;
RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (SSP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL136793; CAB66727.1; -.
DR HSSP; Q60778; IOY3.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 6.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 6.
DR PROSITE; PS00088; ANK_REPEAT; 4.
DR PROSITE; PS02097; ANK_REPEAT_REGION; 1.
KW ANK repeat; Hypothetical protein.
SQ SEQUENCE 823 AA; 93968 MW; 9D2DD5F52C379B2 CRC64;

Query Match 20.0%; Score 519; DB 2; Length 823;
Best Local Similarity 30.1%; Pred. No. 2.1e-13;
Matches 148; Conservative 85; Mismatches 152; Indels 106; Gaps 11;

Qy 36 QKSVNKALELKNEQTLR-----LRADEILPSESQKQDYESSWDSLSLCT 80
Db 249 QQILEHKKILKKEKSDVGSSESAVSIFHELVRVDSLPAASDDKOLNVATKQCVPEKSEP 308
Qy 81 VSQKDVCLPKATHQEKIDKINGKLES PDNDGFLKAPCRMKVSIPTKALELMDMQTFKAE 140
Db 309 -----LPGSSHEKGNRIVNGQE-----G 327
Qy 141 PPEKPSAFPEAEMQKSVNKALELKNEQTLRADQMPP--SESKQKVEENSWDSSESURE 198
Db 328 PPAKHPSLKPSTEVDEPAVKGAVQKQKQVTLRAEQALPVASEEQRHE-----RS 378
Qy 199 TVSQKDVCPKATHQEKMDKISGLKEDSTSLSKILDVHSCERARELOKDHCEORTGKWE 258
Db 379 EKQPOVKEGNTNKSEKIQLSENICDSTS-----SAAAGRLT 416
Qy 259 QMKKFCVKKLSEAKEIKSOLENOKVKEQELCSVRLTLNQBEKKRRNADILNEKIRE 318

```

```
Db 417 QQRK-----IGKTPQPPF-KKLKEBHC-----TLKQNEEKTNNMLYKKNRE 461
Qy 319 ELGRIBEQHRELEVKQLEQALRIQDIELKSVESNLNOVSHENENYLLHENCMLKKE 378
Db 462 ELERKEQYKKEVEAK-QLEPTVOSLEMSKSTARNTPNRDPHNHEEMKGLMDENCILKAD 520
Qy 379 IAMLKLEIATLKHQVQOEKRYFEDIKILKEKNAELQMTLKEESLTKRASQYSGQLKV 438
Db 521 IAILRQELCTWMDNLEKENYKDKIKVKTNALEKYIKLEEMITEITAFRYQOELND 580
Qy 439 LIAENTMLTSK-LKEKQDKLEIAEIESHHPRLASAVODHDOIVTSRKSPAFHIAQDA 497
Db 591 LKAENTRLNAELLKESKGLADIESYQSKLAAAIKSHSVKTERNLKIALERTRDV 640
Qy 498 CLQRKMNVDVS 508
Db 641 SVQVEMSSAIS 651

RESULT 10
Q81ZM7 PRELIMINARY; PRT; 453 AA.
AC Q81ZM7
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Melanoma-associated antigen (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bruno R., d'Orlando O., Altomonte A., Lamaj E., Maio M., Pucillo C.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF543495; AAN40505.1; -.
FT NON TER
SQ SEQUENCE 453 AA; 53140 MW; 81DF80012122BBCA CRC64;

Query Match 18.3%; Score 475; DB 2; Length 453;
Best Local Similarity 45.3%; Pred. No. 6.6e-12;
Matches 107; Conservative 54; Mismatches 62; Indels 6; Gaps 4;

Qy 275 KEIKSOLEKQKWEQELCSVRLTLNQBEKKRNADILNEKIRELGRIEHQHRELEVK 334
Db 2 BEIKSQLKHEILEKEKLSLRFPAQOEKKERNVEVHQVKREKLRTESQYRIEADVT 61
Qy 335 QOQLEALRIQDIELKSVESNLNOVSHENENYLLHENCMLKKEIATLKHQYQ 394
Db 62 KPKEPKAKSAEVELKTGGNSNQVSETDEKED-LLHENRLMQDEIARLRLEKDTIKQNQL 120
Qy 395 EKENKYFEDIKILKEKNAELQMTLKEESLTKRASQYSGQLKVLIAENTMLTSKL-KEK 453
Db 121 EK--KYLKDFEIVKVRKHEDLQALKRNGETLAKTIACYSGQAAUTDENTTLRSKLEQR 178
Qy 454 QBEKLEIAEIESHHPRLASAVODHDOIVTSRKSPAFHIAQDAC--LQRKMNVDV 507
Db 179 ESRQRLETQMOSYHCLNARNACDHQSHSKEDQLAQGTVDYKCRHLQENLSHV 234

RESULT 11
O60311 PRELIMINARY; PRT; 641 AA.
AC O60311
DT 01-AUG-1998 (TremBLrel. 07, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TremBLrel. 22, Last annotation update)
DE KIAA0565 protein (Fragment).
GN Name=KIAA0565;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```

```
RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RX MEDLINE=98290545; PubMed=9628581;
RA Nagase T., Ishikawa K., Miyajima N., Tanaka A., Kotani H., Nomura N.,
RA Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. IX.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 5:31-39(1998).
DR EMBL; AB011137; BAA25491.2; -.
FT NON TER
SQ SEQUENCE 641 AA; 75777 MW; F11566F2E6A0B9F4 CRC64;

Query Match 16.6%; Score 431.5; DB 2; Length 641;
Best Local Similarity 40.2%; Pred. No. 5.2e-10;
Matches 106; Conservative 43; Mismatches 58; Indels 57; Gaps 4;

Qy 296 RLTLNQBEKKRNADILNEKIRELGRIEHQHRELEVKQLEQALRIQDIELKSVESNL 355
Db 6 RLALKQNEEKKRNADMLYNKDSQELRIKEBECGVKQVKQKLNRLRLVKELRTVRNL 65
Qy 356 -----NOVSHENENYLLHEN 372
Db 66 DLVVQERNDAQQLSEQDARILODQILTSKQKELEMARKKNNSEISHRHQKEDLFED 125
Qy 373 CMLKKEIATLKHQVQOEKRYFEDIKILKEKNAELQMTLKEESLTKRASQY 432
Db 126 CMLQEEIALRLLEIDITIKNQNKQKQKQYFEDIEAVKKNLQKIILKEETLTETILQY 185
Qy 433 SGQLKVLIAENTMLTSKLKE-KQDKLEIAEIESHHPRLASAVODHDOIVTSR----- 484
Db 186 SGQLNLTAEKNILNSELNGKQOELEIMESYRCRLAAAVRDCQSQOTARDLKLDFO 245
Qy 485 KSQEPAFHIAQDACLQRKMNVDVS 508
Db 246 RTQEWVR-----LHDKMKVDMS 263

RESULT 12
Q7K5Q6 PRELIMINARY; PRT; 2006 AA.
ID Q7K5Q6
AC Q7K5Q6;
DT 05-JUL-2004 (TremBLrel. 27, Created)
DT 05-JUL-2004 (TremBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TremBLrel. 27, Last annotation update)
DE Erythrocyte binding protein 3.
GN Name=maebi; falciparum.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22072559; PubMed=12076768; DOI=10.1016/S0166-6851(02)00067-1;
RA Blair P.L., Kappe S.H., Maciel J.E., Balu B., Adams J.H.;
RT "Plasmodium falciparum MAEBL is a unique member of the ebl family.";
RL Mol. Biochem. Parasitol. 122:35-44(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22077637; PubMed=12082132;
RA Michon P., Stevens J.R., Kaneko O., Adams J.H.;
RT "Evolutionary relationships of conserved cysteine-rich motifs in
RT adhesive molecules of malaria parasites.";
RL Mol. Biol. Evol. 19:1128-1142(2002).
RN [3]
RP SEQUENCE FROM N.A.
RA Blair B.L., Kappe S.H.I., Maciel J.E., Balu D.J., Adams B.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY042084; AAQ73469.1; -.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
DR SEQUENCE 2006 AA; 237620 MW; EE75EF42E2BF767D CRC64;
```

Query Match 12.0%; Score 313; DB 2; Length 2006;
 Best Local Similarity 27.3%; Pred. No. 9.4e-05;
 Matches 138; Conservative 94; Mismatches 186; Indels 88; Gaps 22;

QY 19 KAEPPEKSAFPAIEMQKSVNKALELKN--EOTLRADAILPSSKQDYESSWDSSES 76
 DB 1293 KAEKKKKADELKKAEEKK---KADELKKAEEKKKADEVKKAEEKKKADELKKAEEKK 1348
 QY 77 LCETVSQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQMT 136
 DB 1349 KADE-----LKKSEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKK 1393
 QY 137 FKAEPPEKSAFPAIEMQKSVNKALELKN--EOTLRADQMFPSKQKQKVEENSWSDE 194
 DB 1394 -KAEKKKADELKKAEEKK---KADELKKAEEKKKADELKKAEEKKKADELKKAEEKK 1441
 QY 195 SLRTVSQK--DVCVPKATHQKEMDKISGLKEDSTSLSKILDTVHSCERARELOKHCEQ 252
 DB 1442 NLKKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKK 1501
 QY 253 RTGKME--QMKKKFCVLKK-----KLSEAKETKSOLENOKV--KWEQELCSVRL-- 297
 DB 1502 KADELKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKK 1561
 QY 298 --TLNQ-----EKKRNADILNEKIREELGRIBEOHRLKLEVKQLEQALR 342
 DB 1562 AEILKQIEKKRIEVMKLYEEKKMAEQLKKEEEKIK--AEQLKKEEEKKKVQQLKK 1619
 QY 343 IQDIELKSVESNLQVSHTHENENYLLHNCMLKKEIAMLKLEIATLKHQVQEKENKTYFE 402
 DB 1620 KEEEEKKKAQOLKKE-----EENKIKAEQLKKEEEKKKAEL--ELKKE-EEEEKKAE 1671
 QY 403 DIKILKKNALQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSLKLEKQDKKILBAE 462
 DB 1672 QLKKEEEKKKVEQ-LKKKEEEKKAE-----QLKKEEENKIKVQELKKEEEKKAE 1726
 QY 463 IESHHPRLASAVQDHDQIVTSRKSOE 488
 DB 1727 -----ELKKEEEKKKVQQLKKEE 1746

RESULT 13
 Q7K5Q5 PRELIMINARY; PRT; 2019 AA.
 AC Q7K5Q5;
 DT 05-JUL-2004 (TremBLrel. 27, Created)
 DT 05-JUL-2004 (TremBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TremBLrel. 27, Last annotation update)
 DE Erythrocyte binding protein 2.
 GN Name=maeb1;
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22072559; PubMed=12076768; DOI=10.1016/S0166-6851(02)00067-1;
 RA Blair P.L., Kappe S.H., Maciel J.E., Balu B., Adams J.H.;
 RT "Plasmodium falciparum MAEBL is a unique member of the ebl family.";
 RL Mol. Biochem. Parasitol. 122:35-44 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22077637; PubMed=12082132;
 RA Michon P., Stevens J.R., Kaneko O., Adams J.H.;
 RT "Evolutionary relationships of conserved cysteine-rich motifs in
 RT adhesive molecules of malaria parasites.";
 RL Mol. Biol. Evol. 19:1128-1142 (2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Blair B.L., Kappe S.H.I., Maciel J.E., Balu D.J., Adams B.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY042084; AAQ73468.1; -
 DR InterPro; IPR008602; Duffy_binding.
 DR Pfam; PF05424; Duffy_binding; 1.

SQ SEQUENCE 2019 AA; 239410 MW; A626F2C684C08785 CRC64;
 Query Match 12.0%; Score 313; DB 2; Length 2019;
 Best Local Similarity 27.3%; Pred. No. 9.5e-05;
 Matches 138; Conservative 94; Mismatches 186; Indels 88; Gaps 22;

QY 19 KAEPPEKSAFPAIEMQKSVNKALELKN--EOTLRADAILPSSKQDYESSWDSSES 76
 DB 1293 KAEKKKKADELKKAEEKK---KADELKKAEEKKKADEVKKAEEKKKADELKKAEEKK 1348
 QY 77 LCETVSQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQMT 136
 DB 1349 KADE-----LKKSEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKK 1393
 QY 137 FKAEPPEKSAFPAIEMQKSVNKALELKN--EOTLRADQMFPSKQKQKVEENSWSDE 194
 DB 1394 -KAEKKKADELKKAEEKK---KADELKKAEEKKKADELKKAEEKKKADELKKAEEKK 1441
 QY 195 SLRTVSQK--DVCVPKATHQKEMDKISGLKEDSTSLSKILDTVHSCERARELOKHCEQ 252
 DB 1442 NLKKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKK 1501
 QY 253 RTGKME--QMKKKFCVLKK-----KLSEAKETKSOLENOKV--KWEQELCSVRL-- 297
 DB 1502 KADELKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKK 1561
 QY 298 --TLNQ-----EKKRNADILNEKIREELGRIBEOHRLKLEVKQLEQALR 342
 DB 1562 AEILKQIEKKRIEVMKLYEEKKMAEQLKKEEEKIK--AEQLKKEEEKKKVQQLKK 1619
 QY 343 IQDIELKSVESNLQVSHTHENENYLLHNCMLKKEIAMLKLEIATLKHQVQEKENKTYFE 402
 DB 1620 KEEEEKKKAQOLKKE-----EENKIKAEQLKKEEEKKKAEL--ELKKE-EEEEKKAE 1671
 QY 403 DIKILKKNALQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSLKLEKQDKKILBAE 462
 DB 1672 QLKKEEEKKKVEQ-LKKKEEEKKAE-----QLKKEEENKIKVQELKKEEEKKAE 1726
 QY 463 IESHHPRLASAVQDHDQIVTSRKSOE 488
 DB 1727 -----ELKKEEEKKKVQQLKKEE 1746

RESULT 14
 Q8IHP3 PRELIMINARY; PRT; 2055 AA.
 AC Q8IHP3;
 DT 01-MAR-2003 (TremBLrel. 23, Created)
 DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
 DE MAEBL, putative
 GN ORFNames=PF11_0486;
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2255705; PubMed=12368864; DOI=10.1038/nature01097;
 RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
 RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
 RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
 RA Chan M.S., Nene V., Shallow S.J., Sub B., Peterson J., Angiuoli S.,
 RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
 RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
 RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
 RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
 RA Fraser C.M., Barrell B.;
 RT "Genome sequence of the human malaria parasite Plasmodium
 RT falciparum.";
 RL Nature 419:498-511 (2002).
 DR EMBL; AE014843; AAN36066.1; -
 DR HSSP; P04268; 1IC2.
 DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO:0004872; F:receptor activity; IEA.
 DR GO:0009405; P:pathogenesis; IEA.
 DR InterPro: IPR008602; Duffy binding.
 DR Pfam: PF05424; Duffy binding; 1. 783F48BB7286B1F3 CRC64;
 SQ SEQUENCE 2055 AA; 243211 MW; 783F48BB7286B1F3 CRC64;

Query Match 12.0%; Score 313; DB 2; Length 2055;
 Best Local Similarity 27.3%; Pred. No. 9.6e-05;
 Matches 138; Conservative 94; Mismatches 186; Indels 88; Gaps 22;

QY 19 KAEPPEKPSAFAPEAIEMQKSVNKALELKN--EQTLRADQLPSESQKDYEESSWDS 76
 Db 1293 KAEKKKADLKKAEKKK----KADLKKAEKKKADLKKAEKKKADLKKSEBK 1348
 QY 77 LCETVSQDVCLPKATHQKEIDKINGKLESPDNGFLKAPCRMKVSIPTKALELMDMQT 136
 Db 1349 KADE-----LKKSEKKKADLKKAEKKKADLKK-----KAEKKKADLKK 1393
 QY 137 FKAEPPEKPSAFAPEAIEMQKSVNKALELKN--EQTLRADQMFPSQKDYEESSWDS 194
 Db 1394 -KAEKKKADLKKAEKKK----KADLKKAEKKKADLKKAEKKKADLKKAEKK 1441
 QY 195 SLRTVSQK--DVCVPKATHQKEMDKISGKLESDTSLSKILDVHSCRRARELOKHCEQ 252
 Db 1442 NLKAEKKKADLKKAEKKKADLKKAEKKKADLKKAEKKKADLKKAEKKKADLKKAEKK 1501
 QY 253 RTGKME--QMKKFCVLKK-----KLSAKETKSQLENQKV---KWEQLCSVRL-- 297
 Db 1502 KADELKKAEKKKADLKKAEKKKADLKKAEKKKADLKKAEKKKADLKKAEKKKADLKKAEKK 1561
 QY 298 --TLNQ-----EKKERNADILNEKIREELGRIEEEOHRELEVKQLEQALR 342
 Db 1562 AEILKQIEKKRIEVMKLYEEKKKAEQKAEKKKADLKKAEKKKADLKKAEKKKADLKKAEKK 1619
 QY 343 IQDIELKSVEINLQVSHTHENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKPYE 402
 Db 1620 KEESEKKKAEQKKKE-----EENKIKAEQKLEEEEEKKAE--ELAKE--EESKKKAE 1671
 QY 403 DIKILKKNAELOMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSLKKEKQDKELRAE 462
 Db 1672 QLKKEEEKKKVEQ--LKKKEEEKKAE-----OLKKEEENKIKVEQLKKEEEKKKAE 1726
 QY 463 IESHHPRLASAVQDHDQIVTSRKSOE 488
 Db 1727 -----ELKKEEEKKKVQQLKKEE 1746

RESULT 15
 Q8T5C7 PRELIMINARY; PRT; 2055 AA.
 AC Q8T5C7;
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
 DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
 DE Erythrocyte binding protein 1 (Chimeric erythrocyte-binding protein
 DE MAEBL).
 GN Name=maeb1; Synonyms=MAEBL;
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22072559; PubMed=12076768; DOI=10.1016/S0166-6851(02)00067-1;
 RA Blair P.L., Kappe S.H., Maciel J.E., Balu B., Adams J.H.;
 RT "Plasmodium falciparum MAEBL is a unique member of the ebl family.";
 RL Mol. Biochem. Parasitol. 122:35-44 (2002).
 RP [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22077637; PubMed=12082132;
 RA Michon P., Stevens J.R., Kaneko O., Adams J.H.;
 RT "Evolutionary relationships of conserved cysteine-rich motifs in
 RT adhesive molecules of malaria parasites.";
 RL Mol. Biol. Evol. 19:1128-1142 (2002).

RN SEQUENCE FROM N.A.
 RP Blair B.L., Kappe S.H., Maciel J.E., Balu B., Adams J.E., Adams B.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22155635; PubMed=12165387; DOI=10.1016/S0166-6851(02)00123-8;
 RA Ghai M., Dutta S., Hall T., Freilich D., Ockenhouse C.;
 RT "Identification, expression, and functional characterization of MAEBL,
 RT a sporozoite and asexual blood stage chimeric erythrocyte-binding
 RT protein of Plasmodium falciparum.";
 RL Mol. Biochem. Parasitol. 123:35-45 (2002).
 DR EMBL; AY042084; AAL10509.1; -;
 DR EMBL; AF400002; AAM90625.1; -;
 DR HSSP; P04268; 1IC2.
 DR GO:0016021; C:integral to membrane; IEA.
 DR GO:0004872; P:receptor activity; IEA.
 DR GO:0009405; P:pathogenesis; IEA.
 DR InterPro: IPR008602; Duffy binding.
 DR Pfam: PF05424; Duffy binding; 1.
 SQ SEQUENCE 2055 AA; 243211 MW; 783F48BB7286B1F3 CRC64;

Query Match 12.0%; Score 313; DB 2; Length 2055;
 Best Local Similarity 27.3%; Pred. No. 9.6e-05;
 Matches 138; Conservative 94; Mismatches 186; Indels 88; Gaps 22;

QY 19 KAEPPEKPSAFAPEAIEMQKSVNKALELKN--EQTLRADQLPSESQKDYEESSWDS 76
 Db 1293 KAEKKKADLKKAEKKK----KADLKKAEKKKADLKKAEKKKADLKKSEBK 1348
 QY 77 LCETVSQDVCLPKATHQKEIDKINGKLESPDNGFLKAPCRMKVSIPTKALELMDMQT 136
 Db 1349 KADE-----LKKSEKKKADLKKAEKKKADLKK-----KAEKKKADLKK 1393
 QY 137 FKAEPPEKPSAFAPEAIEMQKSVNKALELKN--EQTLRADQMFPSQKDYEESSWDS 194
 Db 1394 -KAEKKKADLKKAEKKK----KADLKKAEKKKADLKKAEKKKADLKKAEKK 1441
 QY 195 SLRTVSQK--DVCVPKATHQKEMDKISGKLESDTSLSKILDVHSCRRARELOKHCEQ 252
 Db 1442 NLKAEKKKADLKKAEKKKADLKKAEKKKADLKKAEKKKADLKKAEKKKADLKKAEKK 1501
 QY 253 RTGKME--QMKKFCVLKK-----KLSAKETKSQLENQKV---KWEQLCSVRL-- 297
 Db 1502 KADELKKAEKKKADLKKAEKKKADLKKAEKKKADLKKAEKKKADLKKAEKKKADLKKAEKK 1561
 QY 298 --TLNQ-----EKKERNADILNEKIREELGRIEEEOHRELEVKQLEQALR 342
 Db 1562 AEILKQIEKKRIEVMKLYEEKKKAEQKAEKKKADLKKAEKKKADLKKAEKKKADLKKAEKK 1619
 QY 343 IQDIELKSVEINLQVSHTHENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKPYE 402
 Db 1620 KEESEKKKAEQKKKE-----EENKIKAEQKLEEEEEKKAE--ELAKE--EESKKKAE 1671
 QY 403 DIKILKKNAELOMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSLKKEKQDKELRAE 462
 Db 1672 QLKKEEEKKKVEQ--LKKKEEEKKAE-----OLKKEEENKIKVEQLKKEEEKKKAE 1726
 QY 463 IESHHPRLASAVQDHDQIVTSRKSOE 488
 Db 1727 -----ELKKEEEKKKVQQLKKEE 1746

Search completed: February 19, 2005, 02:10:37
 Job time : 66.9274 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:56:15 ; Search time 14.4804 Seconds
(without alignments)
3402.035 Million cell updates/sec

Title: US-09-602-362E-16
Perfect score: 2599
Sequence: 1 MKVSIPTKALELMDMTFKA.....IAGDACLQRXNVDVSTDI 512
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	265	10.2	2954	T14156	kinesin-related pr
2	264	10.2	1738	T14867	interaptin - slime
3	261.5	10.1	2116	A26655	myosin heavy chain
4	255	9.8	2017	A36014	myosin heavy chain
5	255	9.8	2057	S61477	myosin II heavy ch
6	250	9.6	944	S26710	spindle pole body
7	247.5	9.5	2253	T30336	nuclear/mitotic ap
8	244.5	9.4	1979	C71622	hypothetical prote
9	241.5	9.3	1313	F96673	hypothetical prote
10	240.5	9.3	946	S28061	SCP1 protein - rat
11	239.5	9.2	1827	T16270	hypothetical prote
12	239	9.2	1790	S67593	transport protein
13	239	9.2	3259	A56539	giantin - human
14	238	9.2	992	T46337	hypothetical prote
15	238	9.2	3225	I52300	giantin - human
16	237.5	9.1	993	S49461	synaptosomal compl
17	236	9.1	1690	T13030	microtubule bindin
18	235.5	9.1	1642	T08880	NMDA receptor-bind
19	235.5	9.1	2245	T18278	myosin heavy chain
20	232	8.9	544	I36911	involucrin L - dou
21	232	8.9	1388	T08633	serine/threonine-s
22	230.5	8.9	1432	B85431	trichohyalin like
23	230.5	8.9	1939	A46762	myosin alpha heavy
24	230	8.8	2442	T08621	centrosome associa
25	230	8.8	2663	S28261	centromere protein
26	229	8.8	1164	T24806	hypothetical prote
27	229	8.8	1875	T38173	myosin-like protei
28	229	8.8	1979	S03166	myosin heavy chain
29	228.5	8.8	978	A70387	conserved hypothet

30 228 8.8 1392 2 A43336 microtubule-vesicl
31 228 8.8 3488 2 T34418 hypothetical prote
32 227.5 8.8 1961 1 A61231 myosin heavy chain
33 227 8.7 2139 2 T18296 myosin heavy chain
34 226.5 8.7 1940 1 S04090 myosin heavy chain
35 226 8.7 1427 2 S22655 restin - human
36 226 8.7 1939 2 T18372 repeat organellar
37 225.5 8.7 1269 2 F94730 probable myosin he
38 225 8.7 1388 2 S74245 serine/threonine-s
39 224.5 8.6 585 1 A24168 involucrin - human
40 224.5 8.6 1935 1 A37102 myosin beta heavy
41 224.5 8.6 1938 1 S06005 myosin alpha heavy
42 224.5 8.6 1938 2 I49464 alpha cardiac myos
43 224.5 8.6 1939 2 I48175 myosin heavy chain
44 224 8.6 853 2 T51505 hypothetical prote
45 224 8.6 1085 2 F96712 hypothetical prote

ALIGNMENTS

RESULT 1
T14156
kinesin-related protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T14156
R:Wood, K.W.; Sakowicz, R.; Goldstein, L.S.; Cleveland, D.W.
Cell 91, 357-366, 1997
A:Title: CENP-E is a plus end-directed kinetochore motor required for metaphase chromosome
A:Reference number: Z17893; MUID:98028574; PMID:9363944
A:Accession: T14156
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2954 <WOO>
A:Cross-references: UNIPROT:O42263; EMBL:AF027728; NID:G2586070; PID:G2586071; PIDN:AAC6
C:Genetics:
A:Gene: XENP-E
C:Superfamily: centromere protein E; kinesin motor domain homology

Query Match	10.2%	Score	265;	DB	2;	Length	2954;
Best Local Similarity	21.8%	Pred. No.	0.00033;				
Matches	106;	Conservative	108;	Mismatches	172;	Indels	100;
Gaps	15;						
Qy	5	IPTKALELMDMTFKABPPE--KPSAFEPAPEMOKSVPNKALELKNEOTLRADELTPSES	62				
Db	1888	IETLSLSLKEKE-FALEQAQKADAAKTIIDITEKISNIEEQLLOQATNLKETLYERES	1946				
Qy	63	KQDYBESSWDSLSCTVSQKDVCLPKATHQKE-----IDKINGKLEESPDNDFG	113				
Db	1947	LIQCKEQALNTEHLRETLLKSKDLALGMEQERDEAANKVIALTEKMSLSEQINENVT	2006				
Qy	114	LKAPCRMKVSIPTKALELMDMTFKABPPEKPSAFEPAPEMOKSVPNKALELKNEOTLRA	173				
Db	2007	LKEGEGEK-----ETFLQRPSSQKQSSSQMEELRESLTKDLQ-----	2045				
Qy	174	QMPFPSSKQKVBENSWSDE-----SLRETVSQKDVCPKATHQKEMDKISGKLEDS	226				
Db	2046	-----FEAEKEISEATNNEIKNLTAKISSLEEILQNASILNEAVSERENLRHS-QQQLV	2098				
Qy	227	TSLSKILDTVHSCRRARELQKHCEQRTGRMEQMKKFCVLKKKLSEAKEIKSQLENQKV	286				
Db	2099	SELQSLSTLKSRDHFAQSKREKDEAVNKIASLAEEIKILTKEWDFRDSKESLQEQSS	2158				
Qy	287	KWQELCSVRLTNQEEKRRNADILNEKIHREELGRTE--OHRKELEVQKQLEQALRIQ	344				
Db	2159	HLSELCTYKTELQMLKQKQED---INNKLAEKVEVDLLOH-----	2198				
Qy	345	DIELKSVESNLNQVSHNENYLLHENC-----MLKKEIAMLKLEIATLKHOYQEKENKY	400				
Db	2199	---LSSKEQJQJQMLNELNKLNYELCEKXMDIMEKISVRL-----MQNEPQEQEEDV	2251				
Qy	401	FEDIKILKEKNABLQMTLKLKESLTKRASQYSGQLKVLIAENTMTLSKLKEKQDKEILE	460				

Db 2252 AERMDILSRNQEIQ-----ELMEKISAVYSEQ-----HTLLSLSSSE-----LQ 2291
QY 461 AEIESH 466
Db 2292 KETEAH 2297
RESULT 2
T14867
interactin - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T14867
R:Rivero, F.J.; Kuspa, A.; Brokamp, R.; Matzner, M.; Noegel, A.A.
J. Cell Biol. 142, 735-750, 1998
A:Title: Interactin, an actin-binding protein of the alpha-actinin superfamily in Dictyo
te.
A:Reference number: Z18248; MUID:98365468; PMID:9700162
A:Accession: T14867
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1738 <RIV>
A:Cross-references: UNIPROT:O76329; EMBL:AF057019; NID:G3549260; PID:G3549261; PIDN:AAC3
C:Genetics:
A:Gene: abpD
A:Introns: 173/2; 1680/1
Query Match 10.2%; Score 264; DB 2; Length 1738;
Best Local Similarity 22.1%; Pred. No. 0.0002;
Matches 133; Conservative 120; Mismatches 203; Indels 146; Gaps 25;
QY 10 LELMDQTPKABPPKPSAFEP-----AIEMQSVNPKALEKNEQTLRADE 56
Db 585 LQLQDQDIK-----EKEFOFEKQLLSQIDSITTNIOEQYQDFNQLQOEFNFTQOTLNOQE 640
QY 57 I--LPSEKQ--KDYEESSWD-----SESLCETVSQKDVCLPKATHQKE---ID 98
Db 641 THRLTQOLQIINTDYNEKOTQLOSEIKQNTQINEQLKQSEKDEIKELSLNQOQQODE 700
QY 99 KINGLEESPDNDGFLKPCRMKVSIPTRKALE-----LMDMTFFKABPPKPSA 147
Db 701 KINLLLEIKEDCLIE-----RINQQLLENIDLSKYQQLLEFENFK-----744
QY 148 FEPAIEMQSVNPKALEKNEQTLRADQMP-----SESKQKVEENSW 191
Db 745 ----LNSKSEKENQLNELQSKODERFNQDNDEKLEKEKQLOQIEDEFNQYKQOQJSSN 800
QY 192 DSESLRETYSQKDVCPKATHQKEMD--KISGLEDSTSLKILDTVHSCERARELQKH 249
Db 801 IDQOLQSTIE-----LSELKEQELNDSKLEKEKQLOQLOQEFFDLN-----EKNQKH 851
QY 250 CEQRTGKMEQMKKFCVLKKLSEAKEIKSLENQKVK-----EOLCSVRLT 298
Db 852 QDQ----LELLEKQLOQOEYDQNETNQSIENQNLNOQNLINKENLNEKEQELLKLQ 907
QY 299 LQNEEKRR-----NADIINEK-----IREELGRLEEGRKELEVKQOLEQA 340
Db 908 LNOQIEKIQDQOEFKQNSINTELNVNKNKELIQDQDYDQLKQNRN--DEKEND 964
QY 341 LRQDTELKSVENLQVSHTHEN---ENYLLH---ENCMLKKEIAMLKLEIATLKHQYQ 394
Db 965 LIEKENQLKSIQNELNQLNLIKESDHKEQOLKQOSIENDLIEKENIQLOLQ-SOLNEQRQ 1023
QY 395 EKENKYPE---DITKILKEKNAELQMTLKKEESLTKRASQYSQOLKILTA-----ENTMLT 447
Db 1024 QOSNOLSEKQDQNLNIEKNQFQKQEQOLKQOSIENDLFEKENIQIQLOLSQNLNEQRQOOS 1083
QY 448 SKLKEK--QDKTEILA--EIESHPRLASAVQDHDQIVTSKSPAPAFHAGDACLQKWN 504
Db 1084 NQSEKQDQNLNIEKNESQKQOLKQOSIENDLIEKENIQIQLOLNEQRLQSEVS 1143
QY 505 VD 506

Db 1144 ID 1145
RESULT 3
A26655
myosin heavy chain [similarity] - slime mold (Dictyostelium discoideum)
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Dictyostelium discoideum
C>Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 09-Jul-2004
C:Accession: A26655; A24728; S00250
R:Warrick, H.M.; De Lozanne, A.; Leinwand, L.A.; Spudich, J.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 9433-9437, 1986
A:Title: Conserved protein domains in a myosin heavy chain gene from Dictyostelium disco
A:Reference number: A26655; MUID:87092266; PMID:3540939
A:Accession: A26655
A:Molecule type: DNA
A:Residues: 1-2116 <WAR>
A:Cross-references: UNIPROT:P08799; GB:M14628; GB:M11938; NID:G167834; PIDN:AAA33227.1;
R:DeLozanne, A.; Lewis, M.; Spudich, J.A.; Leinwand, L.A.
Proc. Natl. Acad. Sci. U.S.A. 82, 6807-6810, 1985
A:Reference number: A24728; MUID:86016788; PMID:3901008
A:Accession: A24728
A:Molecule type: mRNA
A:Residues: 2035-2116
R:Wagle, G.; Noegel, A.; Scheel, J.; Gerisch, G.
FEBS Lett. 227, 71-75, 1988
A:Title: Phosphorylation of threonine residues on cloned fragments of the Dictyostelium
A:Reference number: S00250; MUID:88112226; PMID:2828113
A:Accession: S00250
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1734-1893 <WAG>
C:Comment: The rod domain is highly periodic, containing a pattern of 7-residue repeats
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; coiled coil; hydrolase; nucleotide binding; P-loop; phos
F:1-818/Domain: globular head <HEAD>
F:89-747/Domain: myosin motor domain homology <MMOT>
F:179-186/Region: nucleotide-binding motif A (P-loop)
F:819-2116/Domain: alpha-helical rod <ROD>
Query Match 10.1%; Score 261.5; DB 2; Length 2116;
Best Local Similarity 22.0%; Pred. No. 0.00032;
Matches 135; Conservative 107; Mismatches 212; Indels 161; Gaps 24;
QY 8 KALELMDQTPKABPPKPSAFAPAIEQMSVNPKNAL-----ELKNEQTLRADEILPSESK 63
Db 894 RELRIR-VEDMESLDEKLALENLQKRSVEKVRDLEELQEQKLRN---TLEKL 948
QY 64 QKDYEESSWD-----SESLCETVSQKDVCLPKATHQKEIDKINGKL-BESPNDGFLKAPC 118
Db 949 KKYEELLEEMKRVNDGQSDTISRLEKI--KDELQKEVEELTESFSESKDKGVLEKTRV 1006
QY 119 RMKVSIPTKALELMDMTFFKAE-----PPEKPSAFAPAIEQMSVNPKNAL-- 163
Db 1007 RLQSELDLTVRLDSETKDKSELLRQKKLEEBELKQVQEAALAAETAAKLAQEAANKLQ 1066
QY 164 -----ELKNEQTLR----- 172
Db 1067 EYTELNKFNSEVTARSNVESKKTLESQLVAVNNELDEENKRDALKKKKALDAMLEE 1126
QY 173 -ADQMFPSESQK-----KVENSWDSLSRETYSQKDVCP-----KATHQEMDKISG 221
Db 1127 MKDQLESTGGEKSKSYDLKVKQES--DMEALRNQISELOSTIAKLEIKSTLSEGEVARIQG 1185
QY 222 KLEDSTSLKILDTVHSCERARELQKHCEORTGME--QMKKFCVLKKLSEAKEIKSQ 280
Db 1186 ELE-----AEQLAKSNVEKQKKVLDLEDK-----SAQLAEETAQA 1224
QY 281 LENQKVKWEQELCSVRLTNQEEERKRNRADILNEKIREBELGRIBEQHKELEVKQOLEQA 340
Db 1225 LDKLKKLEQSELSEVOTQLSEANNKVNNSDSTNKHLETSPNNL-----KLELEAEQAKQA 1280

Qy 341 LRIODIELKSVESNLNQVSHTHENENYLLHNCMLKKEIAMLKUEIATLKHQOE----- 395
Db 1281 LEKKRLGL---ESELKHVNEQLEBEKKQKESN---EKRKVDLEKEVSELKQIEEVAASK 1334
Qy 396 -----KEN-----KYFEDIKILKEKNAELOMTLKLKEESTLKEASQYSQLKVL 440
Db 1335 KAVTEAKNKKSELDDEIKRQYADVSSKDSVQKLTLOAKNEELRNNTAEAEAGOLD--- 1391
Qy 441 AENTMLTSLKLEKODKELEA-----ETESHHPRLASAVQDHQDQVTSRKSQ-SPAFHIA 494
Db 1392 ----RAERSKKAEFDLEAEAVKNELEETAKKVKRAEKAMKAETDYRSTKSELDDAKNVS 1446
Qy 495 GDACLO-RKNQDVVS 508
Db 1447 SEQYVQIKRLNEELS 1461

RESULT 4
A36014
myosin heavy chain, nonmuscle - fruit fly (Drosophila melanogaster)
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Drosophila melanogaster
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: A36014; B36014
R:Ketchum, A.S.; Stewart, C.T.; Stewart, M.; Kiehart, D.P.
Proc. Natl. Acad. Sci. U.S.A. 87, 6316-6320, 1990
A:Title: Complete sequence of the Drosophila nonmuscle myosin heavy-chain transcript: cDNA
A:Reference number: A36014; MUID:90349606; PMID:2117279
A:Accession: A36014
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-2017 <RET>
A:Cross-references: UNIPROT:Q99323; GB:M35012
C:Genetics:
A:Gene: FlyBase:zip
A:Cross-references: FlyBase:FBgn0005634
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; alternative splicing; ATP; coiled coil; hydrolase; methylated
F:1-2017/Product: myosin heavy chain, form I #status predicted <MA1>
F:46-2017/Product: myosin heavy chain, form II #status predicted <MA2>
F:135-815/Domain: myosin motor domain homology <MMOT>
F:225-232/Region: nucleotide-binding motif A (P-loop)
F:598-631/Region: actin binding #status predicted
F:705-727/Region: actin binding #status predicted
F:888-2017/Domain: coiled coil #status predicted
F:888-1328/Region: S2
F:1329-2017/Region: light meromyosin
F:176/Modified site: N6,N6-trimethyllysine (Lys) #status predicted
F:231/Binding site: ATP (Lys) #status predicted
F:745,755/Active site: Cys #status predicted

Query Match 9.8%; Score 255; DB 1; Length 2017;
Best Local Similarity 22.2%; Pred. No. 0.00059;
Matches 122; Conservative 113; Mismatches 182; Indels 132; Gaps 23;

Qy 14 DMOTFKAPPEKPSAFEPKPSAEMOKSVPNKALELNE-----QTLRADRLPESKQ----- 64
Db 1170 DLEAKA-----ARAKAEKVRDLSELEALKNELLDSDTTAAQQLSKRQELAT 1222
Qy 65 --KDYESSWDSSELCTVSKQVCLPKATQKEIDKINGKLE----- 105
Db 1223 LKKSLEETVNHGVLADMRHK-----HSQELNSINDQLENLRKAKTVLEKAKGTLE 1274
Qy 106 -ESPD-----NDGFLKAPCRMVSPTKALELMDMTQFKAEPPKPSAFE- 149
Db 1275 AENADLATELSVNSRQENRRRK---QAESQIAELQVKLAETARSELQEKTKLQ 1331
Qy 150 -----PAIEMOKSVPNKALELKNQTLRADQMPFSESKQK-----KVEENSWSDES 195
Db 1332 EAENITNQLAEALKASAAVKSANMESQLTEAQQLLEETFTQKLGSLSKLRQIESEKEA 1391
Qy 196 LRETYSQKDVCPKATHQK-----EMDKTSGKLESDTSLSKLDTVHSCERARELOKD 248

Db 1392 LQEQLEEDDEA--KRNVERKLAEVTTQMQEIKKKAEBDADLAKELE-----EGKKRLNKD 1444
Qy 249 HCEQRTGMEOMKIKFCVLLKKKLLSEA-KEIKSQLENOKVKWEQELCSVRLTLNQOEEKRR 307
Db 1445 -----IEALERQVVELTAQNDRLDKSKKIOSELEDATIELEAQ-----RTKVELEKKQK 1495
Qy 308 NAD-----ILNPKIREELGRIE-----EQRKSELEVQKQLEQAL-RTODIE--LKSVES 353
Db 1496 NFDKILAEKAKAISQIAQERDTAREAREKETKTVLSVRELDEAFDKIEDLENKRKTLQN 1555
Qy 354 NLNQVSHTHENENYLLHNCMLKKEIAMLKUEIATLKHQOEKEN--KYFEDIKILKEKN 411
Db 1556 ELDDLANTQGTADKXVHE---LEKAKRALESQIAELKQAQNEELEDLQUTSDAKLRLVN 1612
Qy 412 ----AELQMTLKLKESLTKRASQYSQLKVLIAE-----NTMLTSKLKQKQKBEI 458
Db 1613 MQALRSQFERDLAKEGAEKRRGLVKQLRDLDELDEERKQRTAAVASKKLEGDLKE 1672
Qy 459 LEAIESHH 467
Db 1673 IETTMEHN 1681

RESULT 5
S61477
myosin II heavy chain, non-muscle - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 19-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C:Accession: S61477; S65349
R:Mansfield, S.G.; Al-Shirawi, D.Y.; Ketchum, A.S.; Newbern, E.C.; Kiehart, D.P.
J. Mol. Biol. 255, 98-109, 1996
A:Title: Molecular organization and alternative splicing in zipper, the gene that encodes
A:Reference number: S61477; MUID:96144835; PMID:8568878
A:Accession: S61477
A:Molecule type: DNA
A:Residues: 1-2057 <MAN>
A:Cross-references: UNIPROT:Q94987; EMBL:U35816
R:Mansfield, S.G.; Al-Shirawi, D.Y.; Ketchum, A.S.; Newbern, E.C.; Kiehart, D.P.
submitted to the EMBL Data Library, September 1995
A:Reference number: S65349
A:Accession: S65349
A:Molecule type: DNA
A:Residues: 1-1908, 'NL', 1911-2057 <MAW>
A:Cross-references: EMBL:U35816; NID:g1141789; PIDN:AAB09049.1; PID:g1572481
C:Genetics:
A:Gene: zip
A:Cross-references: FlyBase:FBgn0005634
A:Introns: 10/3; 90/3; 162/3; 215/1; 255/3; 263/3; 303/3; 334/3; 1248/1; 1349/3; 1526/3;
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: alternative splicing; ATP; nucleotide binding; P-loop
F:135-855/Domain: myosin motor domain homology <MMOT>
F:225-232/Region: nucleotide-binding motif A (P-loop)

Query Match 9.8%; Score 255; DB 2; Length 2057;
Best Local Similarity 22.2%; Pred. No. 0.00061;
Matches 122; Conservative 113; Mismatches 182; Indels 132; Gaps 23;

Qy 14 DMOTFKAPPEKPSAFEPKPSAEMOKSVPNKALELNE-----QTLRADRLPESKQ----- 64
Db 1210 DLEAKA-----ARAKAEKVRDLSELEALKNELLDSDTTAAQQLSKRQELAT 1262
Qy 65 --KDYESSWDSSELCTVSKQVCLPKATQKEIDKINGKLE----- 105
Db 1263 LKKSLEETVNHGVLADMRHK-----HSQELNSINDQLENLRKAKTVLEKAKGTLE 1314
Qy 106 -ESPD-----NDGFLKAPCRMVSPTKALELMDMTQFKAEPPKPSAFE- 149
Db 1315 AENADLATELSVNSRQENRRRK---QAESQIAELQVKLAETARSELQEKTKLQ 1371
Qy 150 -----PAIEMOKSVPNKALELKNQTLRADQMPFSESKQK-----KVEENSWSDES 195
Db 1372 EAENITNQLAEALKASAAVKSANMESQLTEAQQLLEETFTQKLGSLSKLRQIESEKEA 1431

RESULT 7

T30336

nuclear/v

C;Species:

C; Date:

C;Access
P:Mercedes

K; Methods
Cell 87.

A:Title:

A; Referen

A; Access

A;Status

A; Molecu

A; Residu

A;Clostr
C:Geneti

A: Gene:

—

Query

Best, L.

Matche

٤

3

DB

QY

7

८८

Qy

3

D'b

QY 154 -----MKSVPNKALELNKNEOTLRADQ-----MPPSESKOKK 185
 DB 1244 LQKLELSWAQOEKETEALQAKELFKHVKQVELESQTSFTDSGAEALLYLSEAQROQA 1303
 QY 186 VEENSWDSLESRETYSQKDVCPKATHQKEMDKISGLKEDSTSL-----KILDTVHSC 239
 DB 1304 LTKAEQAEQYQKEIEMKNKVENSL--QAEIKILSSKVTYNEEVSDFEORILLKETSXA 1361
 QY 240 ERARELOKDHCEQRTGRKMEQMKKPCV-----LKKKLESEAKE 276
 DB 1362 KLEEKMQKHLMELEASFKELLEKNCAIDCLTTEAQNKGADQOQMAVDSLQKLSKAE 1421
 QY 277 IKSQLENQKVKW-----EQELSVRLTL--NQE-----PEKERNADILNEK 315
 DB 1422 TNHTLQOEIQAWQKNCAEKEQOISLQONLKSNOSSLBEFASLKHYSQEIITAEARDLMOEK 1481
 QY 316 IREEL--GRIEEOHRKELEVKQ-----LEQALRIQDIELKSVE-----NLNOVSH 361
 DB 1482 HQBELLSHKKLTERRQAELEKAKEDMTIIVLLKEKLNQELQHKFQSENSYSLTQISHL 1541
 QY 362 HENENYLLHENCMLK-----KEIAMLKLEIATLKHQYQ-----KENKYPED 403
 DB 1542 QOVNSQLLQANGSLSQISDOGAKKLESEMSTLKEQHKEMKTLRLQYEKTLREGNKQVOE 1601
 QY 404 I-----KILKEKA--BLQWTLKLEESLTKRASOYSGQLKVLIAENTM 445
 DB 1602 TSLOLETVTSKYDVHYSKVLQDKQTFQEKQORLLQVOELNKLQSL-----QEKTI 1652
 QY 446 LTSKLKEQDKEILLEAEATESHPR--LASAVODHDQIVTSRKSQEPAPFIAGDACLQK 502
 DB 1653 RSQOQKLQREGTHEEADKSHKRVLELESQLEQQTQAVEHYKAQMEKAKVHYDA--KKK 1710
 QY 503 MNVDVS 508
 DB 1711 QNQEELS 1716

RESULT 8
 C71622
 hypothetical protein PFB0145c - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004
 C:Accession: C71622
 R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
 :; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
 Science 282, 1126-1132, 1998
 A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
 A:Reference number: A71600; MUID:99021743; PMID:9804551
 A:Accession: C71622
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1979 <GAR>
 A:Cross-references: UNIPROT:O96133; GB:AE001375; GB:AE001362; NID:g3845105; PIDN:AAC7181
 A:Experimental source: clone 3D7
 C:Genetics:
 A:Gene: PFB0145c

Query Match 9.4%; Score 244.5; DB 2; Length 1979;
 Best Local Similarity 23.2%; Pred. No. 0.0017;
 Matches 141; Conservative 92; Mismatches 183; Indels 191; Gaps 30;
 QY 34 EMQKSVNKALELNKNE--QTLR-----ADEILPSEK-----QKDYEE 69
 DB 118 EISKNIHKNELENQLKDTLKSISLSKNKVNYESKIELEKELKVKDKNNDNDYEN 177
 QY 70 SSWSESLECVTSQK-DVCLPKAT--HOKED-----KINGK----- 103
 DB 178 KLEKE---DFVQKIDMLNEKLLQKELDINKREKINEKKNIIKKEETFNIEKE 234
 QY 104 -LEESPNDGFLKAPCRMKVSIPTKALELMD---MOTFKAEPPKPSAREPATEMQKSV 158
 DB 235 YLEKNKERE-----TISIEIIDIKHLEKLEIIEKKEDLENLKNLLSK 280

QY 159 PNKALELK-----NEQTLRA--DOMPPSESKOKKVE-----ENSWDSSES 195
 DB 281 ENVLKELKGCVKCKNETINSNDNIIIEKKYKYLLEVELEKNKQIDLLINKOEKEKEK 340
 QY 196 LRETYSQKDVCPKATHQKEMDKISGLKEDSTSLKILDTVHSCB--RARELOKD-----H 249
 DB 341 EREKEKE-----KEKEKYDTLIKELKDEK--ISILEKVSIIKVRWDIEKREHNFJH 393
 QY 250 CEQRTGRKMEQMKKPCF-----VLKKLSSEAKIKSQLENQKVKWEOELC-----SVRL 297
 DB 394 MED--QLKDLKSVFVKNNNKLVKYKCEIKNLKTELE---KKEKELADIENVSKEINK 446
 QY 298 TLNQEERKRNADILNEKIRE-----LGRIEHQHREL--EVQK-----QLEQ 339
 DB 447 LINQNEKEQILAFNKNHKEEIHGLKEELKESVKITKIETQELQEMVDIIKQKELDQLE 506
 QY 340 ALRIQ-----DIELKSVESNLNOVSHTHNE-----NYLL 369
 DB 507 KYNQAIRESISELSEKKEKYNQYKNTYEEINNLKEETNKETYNLQNTYTNEINMLN 566
 QY 370 HENCMKKEIAMLKLEIATLK---HOYQEKENKYPEDIKILKEKNAELQ---MTLKJKEE 423
 DB 567 NDIHMLNGNIKTMTQISTLKNQVHLLNEQIDKLNNKGTLSKISELNVQIMDLKBEKD 626
 QY 424 SLTKRASOYSGQLKVL-----AENTMLTSKLKEQDKEIILEAEIESHHPRLASAVQHD 478
 DB 627 FLNNQIVDLSNQIDLLTRKMEKENKMLEQENKYQEMELLRGNIS-----SENILNND 681
 QY 479 QIVTSRK 485
 DB 682 EVCOLK 688

RESULT 9
 F96673
 hypothetical protein F13011.30 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: F96673
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 :; ansen, N.F.; Hughes, B.; Huizak, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
 :; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 :; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 :; ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: F96673
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1313 <STO>
 A:Cross-references: UNIPROT:Q9XIP6; GB:AE005173; NID:g5042434; PIDN:AAD38273.1; GSPDB:GN
 C:Genetics:
 A:Gene: F13011.30
 A:Map position: 1

Query Match 9.3%; Score 241.5; DB 2; Length 1313;
 Best Local Similarity 24.5%; Pred. No. 0.0015;
 Matches 146; Conservative 99; Mismatches 185; Indels 165; Gaps 28;
 QY 32 AIEQKSVNKALELNKNE---QTLRADE---ILPSEKQKDYESSWDSSELCTVTSQKD 85
 DB 579 SMKLESLDKBEELKNVTABISSIREWEGSVLEKIELSKVKSGLVDKFTKLQISITGEA 638
 QY 86 VCLP--KATHQKED-----KINGKLEESPD-----NDGFLKAPCRMKVSIPT 125
 DB 639 BELGREAAHMKQIEELSTANASLVDEATKLQISIVQESDLKEKEAGLYKIEELSVANE 698
 QY 126 TKALELMDMOTFKAE---PEKPSAFEPAIEMQKSVPNKALELNKNEQTLRADQMP--- 178

699 SLADNVTDLSQIVQESKDLKEREVAYLKKIE-ELSVANESLVVDKETHKLOHIDQAEELRG 757
179 -SEKQKQKVEENSWDSRLSTVS-----QXDVCPKATHQKEMDKIS---GKLEDS 226
758 REASHLKKIEELSKENENLVNVMQNIABESKDLREREVAYLKKIDELSTANGTLADN 817
227 TS-----LSKILDTVHSCERAR-----243
818 VTNLQNISEENKELRERETITLKKAEELSELNVDKASKLQTVVQVQENELRERETAYL 877
244 ---ELQKH---CEORT-----GKWEOMKKKFCVLKKLSEAKEIKSOLNOKVWEQ 290
878 KIEELSKLHEILSDQETKLOISNHEELKERETAYLKKIEELSKVQEDLLNK-----EN 933
291 ELCVRLTLNOEEKRRNADILNKEIRBELGRIBEOHRKELEVKQOOLEQALRIODIELKS 350
934 ELHGMV---EIEDLRKSDSLAOKKIEELSNF---NASLLIKENELQAVVCNEBELKS 985
351 VE-SNLNQVSHTHENENYLH-----ENCMKKKEIAMKL-----EATLKHQVQE 395
986 KOVSTLKTIDELSLDKQSLIHKELQAAIVENEKLKAE-AALSQRIBELTNLKTQTLID 1044
396 KENK---VFEDIKILKKNAELOMTLK-----LKEESLTKRASQVSGQLK---437
1045 KQNELQGVPHNEELKAEAS---SLKKIDELHLLEQSWLEKESFQVQTQENLELKTOD 1101
438 VLIANTMTLTKKEKQDKKEILEABIESHHPLASAVQDHDQIVTVSRKQSPAFH 492
1102 ALAAKKIEELSKLKE---SLLEKETEL-KCREAAALE-----KMEFPKSH 1142

RESULT 10
SCPI protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 08-Oct-1999
C:Accession: S28061
R:Neuwissen, R.L.J.; Offenberger, H.H.; Dietrich, A.J.J.; Riesenwijk, A.; van Iersel, M.; HEMO J. 1, 5091-5100, 1992
A:Title: A coiled-coil related protein specific for synapsed regions of meiotic prophase
A:Reference number: S28061; MUID:93099884; PMID:1464329
A:Accession: S28061
A:Molecule type: mRNA
A:Residues: 1-946 <NEU>
A:Cross-references: EMBL:X67805; NID:957212; PIDN:CAA48006.1; PID:957213
C:Genetics:
A:Gene: SCPI
C:Keywords: DNA binding

Query Match 9.3%; Score 240.5; DB 2; Length 946;
Best Local Similarity 22.4%; Pred. No. 0.0011;
Matches 139; Conservative 103; Mismatches 193; Indels 185; Gaps 28;

QY 13 MDMQTPKAEPPKPSAFPAIEQMOKSVPNKALELNEQTLRADEILPSSKQKD-----YE 68
DB 187 LEWHFKLEDEHKEIQLHEE--EYQEVNKK-----ENQVSLLLIQSTEKENKMDLTFLE 240
QY 69 ESSWDSSELCTVSKQDVCLPKATHQKIDKINGKLEESPDNDGFLKPCRMKVSIPTKA 128
DB 241 ESRDKANQLEBKTQLDENLKEKNEK--DHLTSELD-----IKSMQORSMTQK 289
QY 129 LELMDMO---TFKAEPPKPSAFPAIEQMOKSVPNKAL---ELKN-----EQLTRADQM 176
DB 290 TLEEDLQIATYIYQLTTEKEAQNE---ELNKAHTHSLVVTTELKATCTTLEELLR----342
QY 177 FPSSEKQKQVEENSWDSSESIRETVSKQDVCPKATHQKEMDKISG-----KLEDSLSL---229
DB 343 ---TEQORLENNE-DQLKLITMBELQK-----KSELEEMTKFNKNKEVELELKITLAE 392
QY 230 -SKILDTVHSCER-ARELOKHCE-----QRTGKMEQMKKFCVLKKK-----LSBAKEI 277
DB 393 DQKLLDEKKQVEKLAELQKGEQELTFLQTRKEIHDLEVOVTVTKTSEHYLKVQVEEM 452

278 KSQLENQKVKQEOELCSVRLTLNQE-----EKKRNADILNEKIRBE 319
DB 453 KTELEKEKLK-----NIELTANSMLLENKLVQEASDMVLELKKHQEDIIINCKQBE 506
QY 320 --LGRIBEOHRKELEVKQOOLEALR--IQ-----DIELK 349
DB 507 RMLKQIETLEEKEMNLRLDELSVRKEFTQOGDEVCKLDKSEENARSIEYEVVLKKEKQMK 566
QY 350 SVESNLN---QVSHTHENENYLHENCMLKK-----EIAMLKLEIATLKH 391
DB 567 ILENKNNLKKOIKENSKNIBELHQENKALKKSSAENKQNLNAYELKYNKLELELASTQK 626
QY 392 QYQEKENKFFEDIKILK-----EKNABLOMTLKLKEESLTKRASQVSGQLKVLI 440
DB 627 KFEEMINNYQKBEIKKISEEKLGLGEVEKAKATVDEAVKLOKE-IDLRQCHKIAEMVALM 685
QY 441 AENTMTLTKKEKQDKKEI-----LEAEIESHHPLAS-----472
DB 686 EKHKHQYDKIVEERDELGLYKNREQEQSSAKVALETLSNIRNELVSLKKQLEVEKEEK 745
QY 473 ----AVQDHDQIVTVSRKSOE 488
DB 746 EKLMEQENTAILTKDKKK 765

RESULT 11
Tl6270
hypothetical protein F35D11.11 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: Tl6270
R:Fulton, B.
submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid F35D11.
A:Reference number: Zl8487
A:Accession: Tl6270
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1827 <FUL>
A:Cross-references: UNIPROT:Q20042; EMBL:U29381; NID:9868214; PID:9868224; PIDN:AAA68757
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F35D11.11
A:Introns: 76/2; 131/3; 159/3; 185/3; 221/3; 253/3; 320/1; 869/3; 1133/3; 1205/2; 1350/1

Query Match 9.2%; Score 239.5; DB 2; Length 1827;
Best Local Similarity 24.7%; Pred. No. 0.0026;
Matches 123; Conservative 96; Mismatches 177; Indels 101; Gaps 24;

QY 42 KALELNEQTLRA---DEILPSESQKQDYESSWDSSELCT-----VSQKDV 86
DB 448 KEREDSHBEALRSKDTTEERSRFEKESRERKIRRESDELRETQSRSLKGDVNAKMTDLDK 507
QY 87 CLPKATHQKIDKINGKLE---ESPNDNGFLKPCRMKVSIPTKALELMDMOTFKAPPE 143
DB 508 TLRLDMLTEDEKLKKLETEREQADQDLIAECRAK-----LDEMAEKEALRK 558
QY 144 KPSAFPAIEQM-----KSVPNKALELNEQTLRADQMPSPSSKQKVEENSWDSSELRET 199
DB 559 ELAEFOAIITAMEGEGKLNQRFLESKNELNTLQI---ESLNSEVENKNEIRNLMAT 615
QY 200 VSQKDVCPVK---ATHQ---KEMDKISGKLED-STLSKILDTVHSCERARELQK-----247
DB 616 LOEKEVHIQNVRTSSHQLTATYVEEANGIDILKAEITLHEQVN--ERTQIASEANEKYD 673
QY 248 DHCEQRTGKMEQ---KKKFCVLKKLSEAKEIKSOLNOKVKEQELCSVR-----LTL 299
DB 674 DAARKNDALLEVDVATQWEKYEQLKMELEEMNRRGQEKEREADLRALLDDLRGNFKLT-732
QY 300 NOEESKRRNADILNKI-----REELGRTEE-QHRKELEVKQOOLEQALRIQ 344
DB 733 NELKQKGVTVDSNLNEISLSLKEQLNKSKEKBEELRMEELEQKNEAEKKEEYEVKLQLA 792

```

Qy 345 DIELKSVESNLNQVSHTHENENYLLHENCMLKKEIATLTKHQYQKKNFYFDI 404
Db 793 EKDRQGV-NFGKECEARMTLTKEH--MLMEEDQLKVD-----HLHTE-----BEV 838
Qy 405 KILKEK-NAELQMTLKLKEESTKTRASQYSGQLKVLIAENTMLT-----SKLKEK 453
Db 839 ERLEKMRKELE---KLEQNDGDRAESNERNRLSSKNEAVTELOBRVOKLEDVVKKE 895
Qy 454 QDKET-LEAEIESHHPR 469
Db 896 EDKEIALRDLDSHEK 912

RESULT 12
S67593
transport protein USO1 - Yeast (Saccharomyces cerevisiae)
N/Alternate names: protein D2552; protein YDL058w
C/Species: Saccharomyces cerevisiae
C/Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
C/Accession: S67593; A38455; S30782
R/Blocher, H.; Brandt, P.
submitted to the Protein Sequence Database, July 1996
A/Reference number: S67587
A/Accession: S67593
A/Molecule type: DNA
A/Residues: 1-1790 <BLO>
A/Cross-references: UNIPROT:Q07380; EMBL:Z74106; NID:G1431058; PID:G1431059
A/Experimental source: strain S288C
R/Nakajima, H.; Hirata, A.; Ogawa, Y.; Yonehara, T.; Yoda, K.; Yamaeaki, M.
J. Cell Biol. 113, 245-260, 1991
A/Title: A cytoskeleton-related gene, USO1, is required for intracellular protein transp
A/Reference number: A38455; MUID:91185402; PMID:2010462
A/Accession: A38455
A/Molecule type: DNA
A/Residues: 1-389, 'TA', 392-724, 'S', 726-1790 <NAK>
A/Cross-references: GB:X43478; NID:G94777; PID:CAA38253.1; PID:G94778
A/Note: the authors translated the codon ACT for residue 768 as Ile
R/Hostetter, M.K.; Herman, D.J.; Bendel, C.M.; McClellan, M.; Tao, N.; Kendrick, K.E.
submitted to the EMBL Data Library, February 1993
A/Description: An integrin analogue in Saccharomyces cerevisiae.
A/Reference number: S30782
A/Accession: S30782
A/Molecule type: DNA
A/Residues: 71-846, 'E', 848-923, 'K', 925-1252, 'I', 1254-1318, 'V', 1320-1460, 'S', 1462-1580, 'S'
A/Cross-references: EMBL:L03188
C/Genetics:
A/Gene: SGD:USO1; INT1
A/Cross-references: SGD:S0002216; MIPS:YDL058w
A/Map position: 4L
C/Keywords: coiled coil; transmembrane protein
F:326-342/Domain: transmembrane #status predicted <TM1>
F:394-410/Domain: transmembrane #status predicted <TM2>
F:617-633/Domain: transmembrane #status predicted <TM3>

Query Match 9.2%; Score 239; DB 2; Length 1790;
Best Local Similarity 22.6%; Pred. No. 0.0026;
Matches 134; Conservative 102; Mismatches 218; Indels 138; Gaps 23;

Qy 4 SIPT--KALELMDMTQTKAEP---PEKPSAFEPALAEIEMQKSVNPALELKNQETLRADEIL 58
Db 840 SIKTFLEGLFTILSQKKAEDGINKMGKDLFALSGEMQAVEEN----- 882
Qy 59 PSEKQKDYESSWDSSELCTVSKQVCLPKATHQKEIDKINGKLEESPDNDGFLKAPC 118
Db 883 -CKNLQEKDKSNVNHQK--ETKSLKEDIKIAKIT---EIKAINENLEE-----MKIQ 929
Qy 119 ----RMKVSIFPKALELMDMTQTKAEPPEKPSAFEPALAEIEMQKSVN--KALELKNQETLR 172
Db 930 NNLSEKEHISKELVE-----YKSRFQSHDLNVLAKTEKLKSLANNYKDMQAEENSILK 983
Qy 173 ADQMFPESEK-----QKKVEENSWDSSELCTVSKQVCLPKATHQKEIDKINGKLEESPDNDGFLKAPC 208
Db 984 AVEESKNESIQLSNLQNKIDMSQOEKENFOIERGSEIKETKLEQKLTISDLQETKEIIS 1043

```

```

Qy 209 KA-----THOKEMDKISGKLEDSTS-----LSKILDTVHSCERARELOKDHCEORTG 255
Db 1044 KSDSSKDEYEQSIILLKEKLEFATTANDENVNKISELTTKTREELAEALAAVKNLKNLELET 1103
Qy 256 KNEQMKKFCVVLKKLSEAKKIKSOLENOQKVKVQELCSVRLTLNQBSEKERNADILNEK 315
Db 1104 KLETSEKALKKVEKNEEHLKEEKIQLEKEATETKQOLNSLRANLESLEKEHEDLAAQILKK 1163
Qy 316 IREELGRIEEOHRKELE-----VKQLEQALRIQDIELKSVESNLNQVSHTHENENYL 368
Db 1164 YEQIANKERYNEEISQINDEITSTQOENSIIKKNDL---EGEVKAMKSTSEQN 1220
Qy 369 LHENCMLKKEIATLTKHQYQKKNFYFDI 411
Db 1221 -----KKSEIDALNLQIKELKKNETNEASLLSESIKSESETVKIKELQDECNFKEV 1274
Qy 412 AELQMTLKLKE-----ESLTKRASQYSGQLKVLIAENTMLTSLKKEQDKREI 458
Db 1275 SELEDKLKASEDKNSKYLELOKSEKIKKEELADAKTTTELKIQLEKITNL-SKAKEKSESEL 1333
Qy 459 LEAEIESHHPRLASAVODHDOI VTSRKSOEPAFHAGDACLQKRMNVDVST 510
Db 1334 SRLKTSSEER-KNABEQLEKLNKIEIQKQAF-----KERLLNEGSS 1378

RESULT 13
A56539
giantin - human
N/Alternate names: macrogolin
C/Species: Homo sapiens (man)
C/Date: 19-Oct-1995 #sequence_revision 26-Jan-1996 #text_change 09-Jul-2004
C/Accession: A56539; S37536
R/Seelig, H.P.; Schranz, P.; Schroeter, H.; Wiemann, C.; Griffiths, G.; Renz, M.
Mol. Cell. Biol. 14, 2564-2576, 1994
A/Title: Molecular genetic analyses of a 376-kilodalton Golgi complex membrane protein (
A/Reference number: A56539; MUID:9418728; PMID:7511208
A/Accession: A56539
A/Molecule type: mRNA
A/Residues: 1-3259 <SEE>
A/Cross-references: UNIPROT:Q14789; EMBL:X75304; NID:G405714; PID:CAA53052.1; PID:G4057
C/Genetics:
A/Gene: GDB:GOLGB1; GCP; GCP371
A/Cross-references: GDB:454958
A/Map position: 3q13.31-3q13.31
C/Superfamily: giantin
C/Keywords: coiled coil; Golgi apparatus; transmembrane protein
F:3238-3254/Domain: transmembrane #status predicted <TM>

```

```

Query Match 9.2%; Score 239; DB 1; Length 3259;
Best Local Similarity 19.5%; Pred. No. 0.0051;
Matches 114; Conservative 116; Mismatches 235; Indels 120; Gaps 17;

Qy 8 KALELMDMTQTKAEPPEKPSAFEPALAEIEMQKSVNPALEL-----KNEQTLRADEILPS 60
Db 1594 KSSKIAESTEWQEKHEKLEQYELLQSYENVSNRAERIQHVAVRQEKQELYGKLKST 1653
Qy 61 ESKQKDYESSWDSSELCTVSKQVCLPKATHQK--EIDKINGKLEESPDNDGFLKAPC 118
Db 1654 EANKKETEKOLQEAQEMEEKMKRPFKSKQKQKILEBENDRLRAEVHPAGDTAKEC 1713
Qy 119 RMKVSIFPKALELMDMTQTKAEPPEKPSAFEPALAEIEMQKSVNPALELKNQETLRADQMF 178
Db 1714 -METLLSSNASKKEELERVKMEYETLSKKFQSLSEKDSLSEEVQDLKH-----QLED 1765
Qy 179 SESQKKVEENSWDSSELCTVSKQVCLPKATHQKEMDKISGKLESTSLSKILDTVHS 238
Db 1766 NVSRQANLEATEKHDTNVT-----EEGTQSIPTGETEEDSLS--MSTRPT 1810
Qy 239 CERARELOK-----DHCEORTGMEQMKKFCVVLKKLSEAKKIKSOLENOK 285
Db 1811 CSESVPKSAKSNPAVKDFSSHDSINNLYQIDQLKRIAGLEEEKQKNEFSQTLNEK 1870

```


Search completed: February 19, 2005, 02:12:05
Job time : 18.4804 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 02:11:01 ; Search time 48.0894 Seconds
(without alignments)
3484.083 Million cell updates/sec

Title: US-09-602-362E-16

Perfect score: 2599

Sequence: 1 MKVSIPTKALELMDQTFKA.....IAGDACLQRKXNDVDSSTDI 512

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/FCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2599	100.0	512	16	US-10-181-663-16
2	2589	99.6	1341	13	US-10-007-805-565
3	2589	99.6	1341	14	US-10-076-622-565
4	2589	99.6	1341	14	US-10-177-293-334
5	2589	99.6	1341	14	US-10-124-805-565
6	2589	99.6	1341	15	US-10-088-270A-4
7	2589	99.6	1349	13	US-10-007-805-573
8	2589	99.6	1349	14	US-10-076-622-573
9	2589	99.6	1349	14	US-10-124-805-573
10	2578	99.2	1002	9	US-09-604-287A-475
11	2578	99.2	1002	9	US-09-834-759-475
12	2578	99.2	1002	10	US-09-551-621-475
13	2578	99.2	1002	13	US-10-007-805-475

14	2578	99.2	1002	14	US-10-076-622-475	Sequence 475, App
15	2578	99.2	1002	14	US-10-124-805-475	Sequence 475, App
16	2578	99.2	1095	9	US-09-834-759-493	Sequence 493, App
17	2578	99.2	1095	13	US-10-007-805-493	Sequence 493, App
18	2578	99.2	1095	14	US-10-076-622-493	Sequence 493, App
19	2578	99.2	1095	14	US-10-124-805-493	Sequence 493, App
20	2576	99.1	1013	13	US-10-007-805-553	Sequence 553, App
21	2576	99.1	1013	14	US-10-076-622-553	Sequence 553, App
22	2576	99.1	1013	14	US-10-124-805-553	Sequence 553, App
23	2125	81.8	445	9	US-09-825-301-29	Sequence 29, Appl
24	2125	81.8	445	9	US-09-604-287A-473	Sequence 473, App
25	2125	81.8	445	9	US-09-834-759-473	Sequence 473, App
26	2125	81.8	445	10	US-09-551-621-473	Sequence 473, App
27	2125	81.8	445	13	US-10-007-805-473	Sequence 473, App
28	2125	81.8	445	14	US-10-076-622-473	Sequence 473, App
29	2125	81.8	445	14	US-10-124-805-473	Sequence 473, Appl
30	2125	81.8	445	14	US-10-033-527-29	Sequence 29, Appl
31	1781	68.5	1225	14	US-10-177-293-332	Sequence 332, App
32	1775	68.3	1239	13	US-10-007-805-577	Sequence 577, App
33	1775	68.3	1239	14	US-10-076-622-577	Sequence 577, App
34	1775	68.3	1239	14	US-10-124-805-577	Sequence 577, App
35	1542	59.3	1011	16	US-10-408-765A-1557	Sequence 1557, App
36	1513.5	58.2	650	9	US-09-825-301-25	Sequence 25, Appl
37	1513.5	58.2	650	9	US-09-604-287A-469	Sequence 469, App
38	1513.5	58.2	650	9	US-09-834-759-469	Sequence 469, App
39	1513.5	58.2	650	10	US-09-551-621-469	Sequence 469, App
40	1513.5	58.2	650	13	US-10-007-805-469	Sequence 469, App
41	1513.5	58.2	650	14	US-10-076-622-469	Sequence 469, App
42	1513.5	58.2	650	14	US-10-124-805-469	Sequence 469, App
43	1513.5	58.2	650	14	US-10-033-527-25	Sequence 25, Appl
44	1513.5	58.2	743	9	US-09-834-759-494	Sequence 494, App
45	1513.5	58.2	743	13	US-10-007-805-494	Sequence 494, App

ALIGNMENTS

RESULT 1
US-10-181-663-16
; Sequence 16, Application US/10181663
; Publication No. US20040126398A1
; GENERAL INFORMATION:
; APPLICANT: Jager, Dirk
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gure, Ali
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd
; APPLICANT: Chen, Yao-tseng
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigen
; FILE REFERENCE: LUD 5615
; CURRENT APPLICATION NUMBER: US/10/181.663
; CURRENT FILING DATE: 2002-07-19
; PRIOR FILING DATE: 1999-11-30
; NUMBER OF SEQ ID NOS: 19
; SEQ ID NO 16
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-181-663-16

Query Match	100.0%	Score 2599;	DB 16;	Length 512;
Best Local Similarity	100.0%	Pred. No. 1.7e-141;		
Matches 512;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	MKVSIPTKALELMDQTFKA	PPKPSAFPAIE	MQKSVNKALELKNQTLRADEILPS 60
Db	1	MKVSIPTKALELMDQTFKA	PPKPSAFPAIE	MQKSVNKALELKNQTLRADEILPS 60
Qy	61	ESKQDVESWDSSESICETVSQKDVCLPKATHOKEIDKINGKLBESPDNGFLKAPCRM 120		

Db	61	ESKQDYEESSWDSLSLCEVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM	120
Qy	121	KVSIPTKALELMDMTFKAEPPKPSAFEPALIEHQKSVPNKALELKNQETLRADQMPFSE	180
Db	121	KVSIPTKALELMDMTFKAEPPKPSAFEPALIEHQKSVPNKALELKNQETLRADQMPFSE	180
Qy	181	SKQKVEENSWDSLSLCEVTSQKDVCLPKATHQKEMDKISGKLESDSTLSKILDTVHSC	240
Db	181	SKQKVEENSWDSLSLCEVTSQKDVCLPKATHQKEMDKISGKLESDSTLSKILDTVHSC	240
Qy	241	PARELOKDHCEORTGMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN	300
Db	241	PARELOKDHCEORTGMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN	300
Qy	301	QEEKRRNADILNEKIREELGRIEBQHRKELEVQKQLEQALRIQDIELKSVESNLNOVSH	360
Db	301	QEEKRRNADILNEKIREELGRIEBQHRKELEVQKQLEQALRIQDIELKSVESNLNOVSH	360
Qy	361	THEENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL	420
Db	361	THEENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL	420
Qy	421	KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHDQI	480
Db	421	KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHDQI	480
Qy	481	VTSRKSQEPAFHAGDACLQKRNVDVSSDI	512
Db	481	VTSRKSQEPAFHAGDACLQKRNVDVSSDI	512
RESULT 2			
US-10-007-805-565			
; Sequence 565, Application US/10007805			
; Publication No. US20020150581A1			
; GENERAL INFORMATION:			
; APPLICANT: Jiang, Yugu			
; APPLICANT: Dillon, Davin C.			
; APPLICANT: Mitcham, Jennifer L.			
; APPLICANT: Xu, Jiangchun			
; APPLICANT: Harlocker, Susan L.			
; APPLICANT: Hepler, William T.			
; APPLICANT: Henderson, Robert A.			
; APPLICANT: Fanger, Gary R.			
; APPLICANT: Vedwick, Thomas S.			
; APPLICANT: McNeill, Patricia D.			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY			
; FILE REFERENCE: 210121.470C10			
; CURRENT FILING DATE: 2001-12-07			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 565			
; LENGTH: 1341			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-007-805-565			
Query Match 99.6%; Score 2589; DB 13; Length 1341;			
Best Local Similarity 100.0%; Pred. No. 1.9e-140;			
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MKVSIPTKALELMDMTFKAEPPKPSAFEPALIEHQKSVPNKALELKNQETLRADILPS	60
Db	677	MKVSIPTKALELMDMTFKAEPPKPSAFEPALIEHQKSVPNKALELKNQETLRADILPS	736
Qy	61	ESKQDYEESSWDSLSLCEVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM	120
Db	737	ESKQDYEESSWDSLSLCEVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM	796
Qy	121	KVSIPTKALELMDMTFKAEPPKPSAFEPALIEHQKSVPNKALELKNQETLRADQMPFSE	180
Db	797	KVSIPTKALELMDMTFKAEPPKPSAFEPALIEHQKSVPNKALELKNQETLRADQMPFSE	856
Qy	181	SKQKVEENSWDSLSLCEVTSQKDVCLPKATHQKEMDKISGKLESDSTLSKILDTVHSC	240
Db	857	SKQKVEENSWDSLSLCEVTSQKDVCLPKATHQKEMDKISGKLESDSTLSKILDTVHSC	916
Qy	241	PARELOKDHCEORTGMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN	300
Db	917	PARELOKDHCEORTGMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN	976
RESULT 3			
US-10-076-622-565			
; Sequence 565, Application US/10076622			
; Publication No. US20030023036A1			
; GENERAL INFORMATION:			
; APPLICANT: Houghton, Raymond L.			
; APPLICANT: Sleath, Paul R.			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY			
; FILE REFERENCE: 210121.470C11			
; CURRENT FILING DATE: 2002-02-13			
; NUMBER OF SEQ ID NOS: 627			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 565			
; LENGTH: 1341			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-076-622-565			
Query Match 99.6%; Score 2589; DB 14; Length 1341;			
Best Local Similarity 100.0%; Pred. No. 1.9e-140;			
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MKVSIPTKALELMDMTFKAEPPKPSAFEPALIEHQKSVPNKALELKNQETLRADILPS	60
Db	677	MKVSIPTKALELMDMTFKAEPPKPSAFEPALIEHQKSVPNKALELKNQETLRADILPS	736
Qy	61	ESKQDYEESSWDSLSLCEVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM	120
Db	737	ESKQDYEESSWDSLSLCEVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM	796
Qy	121	KVSIPTKALELMDMTFKAEPPKPSAFEPALIEHQKSVPNKALELKNQETLRADQMPFSE	180
Db	797	KVSIPTKALELMDMTFKAEPPKPSAFEPALIEHQKSVPNKALELKNQETLRADQMPFSE	856
Qy	181	SKQKVEENSWDSLSLCEVTSQKDVCLPKATHQKEMDKISGKLESDSTLSKILDTVHSC	240
Db	857	SKQKVEENSWDSLSLCEVTSQKDVCLPKATHQKEMDKISGKLESDSTLSKILDTVHSC	916
Qy	241	PARELOKDHCEORTGMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN	300
Db	917	PARELOKDHCEORTGMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN	976

QY 301 QEEKRRNADILNEKIREELGRIBEOHREKLEVKQOQLRQALRIQDIELKSVESNLNQVSH 360
Db 977 QEEKRRNADILNEKIREELGRIBEOHREKLEVKQOQLRQALRIQDIELKSVESNLNQVSH 1036
QY 361 THENENYLLHENCMLKKEIAMLKLEIATLKHOYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 1037 THENENYLLHENCMLKKEIAMLKLEIATLKHOYQEKENKYPEDIKILKEKNAELQMTLKL 1096
QY 421 KEESLTKRASQYSGOLKVLIAENTMLTSKLEKQDKEILEAEIESHHPRLASAVQDHDQI 480
Db 1097 KEESLTKRASQYSGOLKVLIAENTMLTSKLEKQDKEILEAEIESHHPRLASAVQDHDQI 1156
QY 481 VTSRKSQEPAFIAGDACLQRKMNVDVSS 510
Db 1157 VTSRKSQEPAFIAGDACLQRKMNVDVSS 1186

RESULT 4

US-10-177-293-334
; Sequence 334, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Puzstai, Lajos
; APPLICANT: Mexic, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 334
; LENGTH: 1341
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-334

Query Match 99.6%; Score 2589; DB 14; Length 1341;
Best Local Similarity 100.0%; Pred. No. 1.9e-140;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADEILPS 60
Db 677 MKVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADEILPS 736

QY 61 ESKQDYEESSWDSSESLCETVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 737 ESKQDYEESSWDSSESLCETVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 796
QY 121 KVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADEILPS 180
Db 797 KVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADEILPS 856
QY 181 SKQKVVENSWDSSESLRETTSQKDVCPKATHQKEMDKISGKLEEDSTSLSKILDTVHSCE 240
Db 857 SKQKVVENSWDSSESLRETTSQKDVCPKATHQKEMDKISGKLEEDSTSLSKILDTVHSCE 916
QY 241 RARELQKDHCEQRTGKMEQMKKFCVLKKLSEAKEIKSOLQENQKVKWEQELCSVRLTLN 300
Db 917 RARELQKDHCEQRTGKMEQMKKFCVLKKLSEAKEIKSOLQENQKVKWEQELCSVRLTLN 976
QY 301 QEEKRRNADILNEKIREELGRIBEOHREKLEVKQOQLRQALRIQDIELKSVESNLNQVSH 360
Db 977 QEEKRRNADILNEKIREELGRIBEOHREKLEVKQOQLRQALRIQDIELKSVESNLNQVSH 1036
QY 361 THENENYLLHENCMLKKEIAMLKLEIATLKHOYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 1037 THENENYLLHENCMLKKEIAMLKLEIATLKHOYQEKENKYPEDIKILKEKNAELQMTLKL 1096
QY 421 KEESLTKRASQYSGOLKVLIAENTMLTSKLEKQDKEILEAEIESHHPRLASAVQDHDQI 480
Db 1097 KEESLTKRASQYSGOLKVLIAENTMLTSKLEKQDKEILEAEIESHHPRLASAVQDHDQI 1156
QY 481 VTSRKSQEPAFIAGDACLQRKMNVDVSS 510
Db 1157 VTSRKSQEPAFIAGDACLQRKMNVDVSS 1186

RESULT 5

US-10-124-805-565
; Sequence 565, Application US/10124805
; Publication No. US20030166022A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C12
; CURRENT APPLICATION NUMBER: US/10/124,805
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 565
; LENGTH: 1341
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-124-805-565

Query Match 99.6%; Score 2589; DB 14; Length 1341;
Best Local Similarity 100.0%; Pred. No. 1.9e-140;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADEILPS 60
Db 677 MKVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADEILPS 736

QY 61 ESKQDYEESSWDSSESLCETVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 737 ESKQDYEESSWDSSESLCETVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 796

QY 121 KVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADEILPS 180
Db 797 KVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADEILPS 856

QY 181 SKQKVVENSWDSSESLRETTSQKDVCPKATHQKEMDKISGKLEEDSTSLSKILDTVHSCE 240
Db 857 SKQKVVENSWDSSESLRETTSQKDVCPKATHQKEMDKISGKLEEDSTSLSKILDTVHSCE 916

Qy 241 RARELOKDHCEORTGMEQMKKFCVLKKLSEAKETKSOLENOKVWEQELCSVRLTLN 300
Db 917 RARELOKDHCEORTGMEQMKKFCVLKKLSEAKETKSOLENOKVWEQELCSVRLTLN 976
Qy 301 QEEKRRNADILNEKIREELGRIBEOHRKELEVQKQLEQALRIQDIELKSVESNLNVSH 360
Db 977 QEEKRRNADILNEKIREELGRIBEOHRKELEVQKQLEQALRIQDIELKSVESNLNVSH 1036
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 1037 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 1096
Qy 421 KEESLTTRASQYSGQLKVLIAENTMTLSKLEKODKEILEAEIESHHPRLASAVQDHDQI 480
Db 1097 KEESLTTRASQYSGQLKVLIAENTMTLSKLEKODKEILEAEIESHHPRLASAVQDHDQI 1156
Qy 481 VTSRKSQEPAPHIAGDACLQKQKMNVDVSST 510
Db 1157 VTSRKSQEPAPHIAGDACLQKQKMNVDVSST 1186

RESULT 6

US-10-058-270A-4

; Sequence 4, Application US/10058270A

; Publication No. US20040029114A1

; GENERAL INFORMATION:

; APPLICANT: Mack, David H.

; APPLICANT: Gish, Kurt C.

; APPLICANT: Afar, Daniel

; APPLICANT: Eos Biotechnology, Inc.

; TITLE OF INVENTION: Methods of Diagnosis of Breast Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Breast Cancer

; FILE REFERENCE: 018501-0052100S

; CURRENT APPLICATION NUMBER: US/10/058,270A

; CURRENT FILING DATE: 2002-01-24

; PRIOR APPLICATION NUMBER: US 60/263,965

; PRIOR FILING DATE: 2001-01-24

; PRIOR APPLICATION NUMBER: US 60/265,928

; PRIOR FILING DATE: 2001-02-02

; PRIOR APPLICATION NUMBER: US 09/829,472

; PRIOR FILING DATE: 2001-04-09

; PRIOR APPLICATION NUMBER: US 60/282,698

; PRIOR FILING DATE: 2001-04-09

; PRIOR APPLICATION NUMBER: US 60/288,590

; PRIOR FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/294,443

; PRIOR FILING DATE: 2001-05-29

; NUMBER OF SEQ ID NOS: 141

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 1341

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-058-270A-4

Query Match 99.6%; Score 2589; DB 15; Length 1341;
Best Local Similarity 100.0%; Pred. No. 1.9e-140; Indels 0; Gaps 0;
Matches 510; Conservative 0; Mismatches 0;

Qy 1 MKVSIPTKALELMDMQTFKAEPPEKPSAFEPALIEQKSVPNKALELKNQETLRADIELPS 60
Db 677 MKVSIPTKALELMDMQTFKAEPPEKPSAFEPALIEQKSVPNKALELKNQETLRADIELPS 736
Qy 61 ESKQKDYESSWDSSESLCETVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 737 ESKQKDYESSWDSSESLCETVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 796
Qy 121 KVSIPTKALELMDMQTFKAEPPEKPSAFEPALIEQKSVPNKALELKNQETLRADQMPFSE 180
Db 797 KVSIPTKALELMDMQTFKAEPPEKPSAFEPALIEQKSVPNKALELKNQETLRADQMPFSE 856
Qy 181 SKQKKEVNSWDSSESLRETQVTSQKDVCPKATHQKEMDKISGKLEDSLSKILDTVHSCE 240

Db 857 SKQKKEVNSWDSSESLRETQVTSQKDVCPKATHQKEMDKISGKLEDSLSKILDTVHSCE 916
Qy 241 RARELOKDHCEORTGMEQMKKFCVLKKLSEAKETKSOLENOKVWEQELCSVRLTLN 300
Db 917 RARELOKDHCEORTGMEQMKKFCVLKKLSEAKETKSOLENOKVWEQELCSVRLTLN 976
Qy 301 QEEKRRNADILNEKIREELGRIBEOHRKELEVQKQLEQALRIQDIELKSVESNLNVSH 360
Db 977 QEEKRRNADILNEKIREELGRIBEOHRKELEVQKQLEQALRIQDIELKSVESNLNVSH 1036
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 1037 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 1096
Qy 421 KEESLTTRASQYSGQLKVLIAENTMTLSKLEKODKEILEAEIESHHPRLASAVQDHDQI 480
Db 1097 KEESLTTRASQYSGQLKVLIAENTMTLSKLEKODKEILEAEIESHHPRLASAVQDHDQI 1156
Qy 481 VTSRKSQEPAPHIAGDACLQKQKMNVDVSST 510
Db 1157 VTSRKSQEPAPHIAGDACLQKQKMNVDVSST 1186

RESULT 7

US-10-007-805-573

; Sequence 573, Application US/10007805

; Publication No. US20020150581A1

; GENERAL INFORMATION:

; APPLICANT: Jiang, Yugu

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Xu, Jiangchun

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedwick, Thomas S.

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Durham, Margarita

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER

; FILE REFERENCE: 210121.470C10

; CURRENT APPLICATION NUMBER: US/10/007,805

; CURRENT FILING DATE: 2001-12-07

; NUMBER OF SEQ ID NOS: 593

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 573

; LENGTH: 1349

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-007-805-573

Query Match 99.6%; Score 2589; DB 13; Length 1349;
Best Local Similarity 100.0%; Pred. No. 1.9e-140; Indels 0; Gaps 0;
Matches 510; Conservative 0; Mismatches 0;

Qy 1 MKVSIPTKALELMDMQTFKAEPPEKPSAFEPALIEQKSVPNKALELKNQETLRADIELPS 60
Db 685 MKVSIPTKALELMDMQTFKAEPPEKPSAFEPALIEQKSVPNKALELKNQETLRADIELPS 744
Qy 61 ESKQKDYESSWDSSESLCETVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 745 ESKQKDYESSWDSSESLCETVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 804
Qy 121 KVSIPTKALELMDMQTFKAEPPEKPSAFEPALIEQKSVPNKALELKNQETLRADQMPFSE 180
Db 805 KVSIPTKALELMDMQTFKAEPPEKPSAFEPALIEQKSVPNKALELKNQETLRADQMPFSE 864
Qy 181 SKQKKEVNSWDSSESLRETQVTSQKDVCPKATHQKEMDKISGKLEDSLSKILDTVHSCE 240
Db 865 SKQKKEVNSWDSSESLRETQVTSQKDVCPKATHQKEMDKISGKLEDSLSKILDTVHSCE 924

```
Qy 241 RARELQDHCORTGKMEOMKKKFCVLKKLSEAKIKSQLENOKVKWEQELCSVRLTLN 300
|
|
|
Db 925 RARELQDHCORTGKMEOMKKKFCVLKKLSEAKIKSQLENOKVKWEQELCSVRLTLN 984
|
|
|
Qy 301 QEEKRRNADILNEKIREELGRIEBOHRKELEVQKQLEQALRIODIELKSYESNLNOVSH 360
|
|
|
Db 985 QEEKRRNADILNEKIREELGRIEBOHRKELEVQKQLEQALRIODIELKSYESNLNOVSH 1044
|
|
|
Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
|
|
|
Db 1045 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 1104
|
|
|
Qy 421 KEESLTKRASQYSGOLKVLIAENTMLTSKLKEKQDKEILEAEIESHHPRLASAVODHDOI 480
|
|
|
Db 1105 KEESLTKRASQYSGOLKVLIAENTMLTSKLKEKQDKEILEAEIESHHPRLASAVODHDOI 1164
|
|
|
Qy 481 VTSRKSQEPAPHIAGDACLQRKMNVDVSST 510
|
|
|
Db 1165 VTSRKSQEPAPHIAGDACLQRKMNVDVSST 1194
|
|
|
RESULT 8
US-10-076-622-573
; Sequence 573, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C11
; CURRENT APPLICATION NUMBER: US/10/076,622
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 573
; LENGTH: 1349
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-622-573
```

```
Query Match 99.6%; Score 2589; DB 14; Length 1349;
Best Local Similarity 100.0%; Pred. No. 1.9e-140;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMTFKAEPPEKPSAFEPAPAEIMQKSVPNKALELKNQETLRADIELPS 60
|
|
|
Db 685 MKVSIPTKALELMDMTFKAEPPEKPSAFEPAPAEIMQKSVPNKALELKNQETLRADIELPS 744
|
|
|
Qy 61 ESKQDYESSWDSLSLCTETVSQKDVCLPKATHOKEIDKINGKLEESPDNDGFLKAPCRM 120
|
|
|
Db 745 ESKQDYESSWDSLSLCTETVSQKDVCLPKATHOKEIDKINGKLEESPDNDGFLKAPCRM 804
|
|
|
Qy 121 KVSIPTKALELMDMTFKAEPPEKPSAFEPAPAEIMQKSVPNKALELKNQETLRADIELPS 180
|
|
|
Db 805 KVSIPTKALELMDMTFKAEPPEKPSAFEPAPAEIMQKSVPNKALELKNQETLRADIELPS 864
|
|
|
Qy 181 SKQKVBENSWSLSLRETTSQKDVCPKATHQKEMDKISGKLESDSTSLSLDTVHSCE 240
|
|
|
Db 865 SKQKVBENSWSLSLRETTSQKDVCPKATHQKEMDKISGKLESDSTSLSLDTVHSCE 924
|
|
|
Qy 241 RARELQDHCORTGKMEOMKKKFCVLKKLSEAKIKSQLENOKVKWEQELCSVRLTLN 300
|
|
|
Db 925 RARELQDHCORTGKMEOMKKKFCVLKKLSEAKIKSQLENOKVKWEQELCSVRLTLN 984
|
|
|
Qy 301 QEEKRRNADILNEKIREELGRIEBOHRKELEVQKQLEQALRIODIELKSYESNLNOVSH 360
|
|
|
Db 985 QEEKRRNADILNEKIREELGRIEBOHRKELEVQKQLEQALRIODIELKSYESNLNOVSH 1044
|
|
|
Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
|
|
|
Db 1045 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 1104
|
|
|
Qy 421 KEESLTKRASQYSGOLKVLIAENTMLTSKLKEKQDKEILEAEIESHHPRLASAVODHDOI 480
|
|
|
Db 1105 KEESLTKRASQYSGOLKVLIAENTMLTSKLKEKQDKEILEAEIESHHPRLASAVODHDOI 1164
|
|
|
Qy 481 VTSRKSQEPAPHIAGDACLQRKMNVDVSST 510
|
|
|
Db 1165 VTSRKSQEPAPHIAGDACLQRKMNVDVSST 1194
|
|
|
RESULT 10
```

```
Qy 421 KEESLTKRASQYSGOLKVLIAENTMLTSKLKEKQDKEILEAEIESHHPRLASAVODHDOI 480
|
|
|
Db 1105 KEESLTKRASQYSGOLKVLIAENTMLTSKLKEKQDKEILEAEIESHHPRLASAVODHDOI 1164
|
|
|
Qy 481 VTSRKSQEPAPHIAGDACLQRKMNVDVSST 510
|
|
|
Db 1165 VTSRKSQEPAPHIAGDACLQRKMNVDVSST 1194
|
|
|
RESULT 9
US-10-124-805-573
; Sequence 573, Application US/10124805
; Publication No. US20030166022A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C12
; CURRENT APPLICATION NUMBER: US/10/124,805
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 573
; LENGTH: 1349
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-124-805-573

Query Match 99.6%; Score 2589; DB 14; Length 1349;
Best Local Similarity 100.0%; Pred. No. 1.9e-140;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMTFKAEPPEKPSAFEPAPAEIMQKSVPNKALELKNQETLRADIELPS 60
|
|
|
Db 685 MKVSIPTKALELMDMTFKAEPPEKPSAFEPAPAEIMQKSVPNKALELKNQETLRADIELPS 744
|
|
|
Qy 61 ESKQDYESSWDSLSLCTETVSQKDVCLPKATHOKEIDKINGKLEESPDNDGFLKAPCRM 120
|
|
|
Db 745 ESKQDYESSWDSLSLCTETVSQKDVCLPKATHOKEIDKINGKLEESPDNDGFLKAPCRM 804
|
|
|
Qy 121 KVSIPTKALELMDMTFKAEPPEKPSAFEPAPAEIMQKSVPNKALELKNQETLRADIELPS 180
|
|
|
Db 805 KVSIPTKALELMDMTFKAEPPEKPSAFEPAPAEIMQKSVPNKALELKNQETLRADIELPS 864
|
|
|
Qy 181 SKQKVBENSWSLSLRETTSQKDVCPKATHQKEMDKISGKLESDSTSLSLDTVHSCE 240
|
|
|
Db 865 SKQKVBENSWSLSLRETTSQKDVCPKATHQKEMDKISGKLESDSTSLSLDTVHSCE 924
|
|
|
Qy 241 RARELQDHCORTGKMEOMKKKFCVLKKLSEAKIKSQLENOKVKWEQELCSVRLTLN 300
|
|
|
Db 925 RARELQDHCORTGKMEOMKKKFCVLKKLSEAKIKSQLENOKVKWEQELCSVRLTLN 984
|
|
|
Qy 301 QEEKRRNADILNEKIREELGRIEBOHRKELEVQKQLEQALRIODIELKSYESNLNOVSH 360
|
|
|
Db 985 QEEKRRNADILNEKIREELGRIEBOHRKELEVQKQLEQALRIODIELKSYESNLNOVSH 1044
|
|
|
Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
|
|
|
Db 1045 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 1104
|
|
|
Qy 421 KEESLTKRASQYSGOLKVLIAENTMLTSKLKEKQDKEILEAEIESHHPRLASAVODHDOI 480
|
|
|
Db 1105 KEESLTKRASQYSGOLKVLIAENTMLTSKLKEKQDKEILEAEIESHHPRLASAVODHDOI 1164
|
|
|
Qy 481 VTSRKSQEPAPHIAGDACLQRKMNVDVSST 510
|
|
|
Db 1165 VTSRKSQEPAPHIAGDACLQRKMNVDVSST 1194
|
|
|
RESULT 10
```

US-09-604-287A-475
; Sequence 475, Application US/09604287A
; Patent No. US20020064872A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-604-287A-475

Query Match 99.2%; Score 2578; DB 9; Length 1002;
Best Local Similarity 99.6%; Pred. No. 5.8e-140;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MKVSIPTKALELMDMTFKAEPPEKPSAFEPALMOKSVNKALELKNQETLRADIELPS 60
DB 338 MKVSIPTKALELMDMTFKAEPPEKPSAFEPALMOKSVNKALELKNQETLRADIELPS 397
QY 61 ESKQDYESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
DB 398 ESKQDYESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457
QY 121 KVSIPTKALELMDMTFKAEPPEKPSAFEPALMOKSVNKALELKNQETLRADIELPS 180
DB 458 KVSIPTKALELMDMTFKAEPPEKPSAFEPALMOKSVNKALELKNQETLRADIELPS 517
QY 181 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 240
DB 518 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 577
QY 241 RARELQKHCEQRTGKMEQMKKFCVLKCKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 300
DB 578 RARELQKHCEQRTGKMEQMKKFCVLKCKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 637
QY 301 QEEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVEENLNQVSH 360
DB 638 QEEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVEENLNQVSH 697
QY 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
DB 698 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 757
QY 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLEKQDKKEILEAIESHHPLASAVDHDQI 480
DB 758 KEESLTKRASQYSGQLKVLIAENTMLTSKLEKQDKKEILEAIESHHPLASAVDHDQI 817
QY 481 VTSRKSQBPAPHIAGDACLQKMNVDVSSST 510
DB 818 VTSRKSQBPAPHIAGDACLQKMNVDVSSST 847

RESULT 11
US-09-834-759-475
; Sequence 475, Application US/09834759
; Publication No. US20020085998A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Jie
; APPLICANT: Dillon, Jennifer L.
; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-834-759-475

Query Match 99.2%; Score 2578; DB 9; Length 1002;
Best Local Similarity 99.6%; Pred. No. 5.8e-140;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MKVSIPTKALELMDMTFKAEPPEKPSAFEPALMOKSVNKALELKNQETLRADIELPS 60
DB 338 MKVSIPTKALELMDMTFKAEPPEKPSAFEPALMOKSVNKALELKNQETLRADIELPS 397
QY 61 ESKQDYESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
DB 398 ESKQDYESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457
QY 121 KVSIPTKALELMDMTFKAEPPEKPSAFEPALMOKSVNKALELKNQETLRADIELPS 180
DB 458 KVSIPTKALELMDMTFKAEPPEKPSAFEPALMOKSVNKALELKNQETLRADIELPS 517
QY 181 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 240
DB 518 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 577
QY 241 RARELQKHCEQRTGKMEQMKKFCVLKCKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 300
DB 578 RARELQKHCEQRTGKMEQMKKFCVLKCKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 637
QY 301 QEEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVEENLNQVSH 360
DB 638 QEEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVEENLNQVSH 697
QY 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
DB 698 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 757
QY 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLEKQDKKEILEAIESHHPLASAVDHDQI 480
DB 758 KEESLTKRASQYSGQLKVLIAENTMLTSKLEKQDKKEILEAIESHHPLASAVDHDQI 817
QY 481 VTSRKSQBPAPHIAGDACLQKMNVDVSSST 510
DB 818 VTSRKSQBPAPHIAGDACLQKMNVDVSSST 847

RESULT 12
US-09-551-621-475
; Sequence 475, Application US/09551621
; Publication No. US20030104366A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Jie
; APPLICANT: Dillon, Jennifer L.
; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C5
; CURRENT APPLICATION NUMBER: US/09/551,621
; CURRENT FILING DATE: 2006-04-17
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-551-621-475

Query Match 99.2%; Score 2578; DB 10; Length 1002;
Best Local Similarity 99.6%; Pred. No. 5.8e-140;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADIELPS 60
Db MKVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADIELPS 397
Qy 61 ESKQDYESSWDSLSLCTVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db ESKQDYESSWDSLSLCTVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457
Qy 121 KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADQMPFSE 180
Db KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADQMPFSE 517
Qy 181 SKQKVVENSWSLSRLTQKDVCPKATHQKEMDKISGKLEDSTSLKILDTVHSC 240
Db SKQKVVENSWSLSRLTQKDVCPKATHQKEMDKISGKLEDSTSLKILDTVHSC 577
Qy 241 RARELQKDHCEORTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQLCSVRLTLN 300
Db RARELQKDHCEORTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQLCSVRLTLN 637
Qy 301 QEEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 360
Db QEEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 697
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHOYEKENKYPEDIKILKEKNAELQMTLKL 420
Db THENENYLLHENCMLKKEIAMLKLEIATLKHOYEKENKYPEDIKILKEKNAELQMTLKL 757
Qy 421 KEESLTKRASQVSGQLKVLIAENTMLTSKLKEKQDKKEILEAIESHHPRLASAVQDDHDI 480
Db KEESLTKRASQVSGQLKVLIAENTMLTSKLKEKQDKKEILEAIESHHPRLASAVQDDHDI 717
Qy 481 VTSRKSQEPAPHIAGDAQLQRKMNVDVSSST 510
Db VTSRKSQEPAPHIAGDAQLQRKMNVDVSSST 847

RESULT 13
US-10-007-805-475
; Sequence 475, Application US/10007805
; Publication No. US20020150581A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedvick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margarita
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C10
; CURRENT APPLICATION NUMBER: US/10/007,805
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 593
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: VARIANT
; LOCATION: 310, 429, 522
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-007-805-475

Query Match 99.2%; Score 2578; DB 13; Length 1002;
Best Local Similarity 99.6%; Pred. No. 5.8e-140;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADIELPS 60
Db MKVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADIELPS 397
Qy 61 ESKQDYESSWDSLSLCTVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db ESKQDYESSWDSLSLCTVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457
Qy 121 KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADQMPFSE 180
Db KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADQMPFSE 517
Qy 181 SKQKVVENSWSLSRLTQKDVCPKATHQKEMDKISGKLEDSTSLKILDTVHSC 240
Db SKQKVVENSWSLSRLTQKDVCPKATHQKEMDKISGKLEDSTSLKILDTVHSC 577
Qy 241 RARELQKDHCEORTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQLCSVRLTLN 300
Db RARELQKDHCEORTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQLCSVRLTLN 637
Qy 301 QEEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 360
Db QEEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 697
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHOYEKENKYPEDIKILKEKNAELQMTLKL 420
Db THENENYLLHENCMLKKEIAMLKLEIATLKHOYEKENKYPEDIKILKEKNAELQMTLKL 757
Qy 421 KEESLTKRASQVSGQLKVLIAENTMLTSKLKEKQDKKEILEAIESHHPRLASAVQDDHDI 480
Db KEESLTKRASQVSGQLKVLIAENTMLTSKLKEKQDKKEILEAIESHHPRLASAVQDDHDI 817
Qy 481 VTSRKSQEPAPHIAGDAQLQRKMNVDVSSST 510
Db VTSRKSQEPAPHIAGDAQLQRKMNVDVSSST 847

RESULT 14
US-10-076-622-475
; Sequence 475, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C11
; CURRENT APPLICATION NUMBER: US/10/076.622

FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 310, 429, 522
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-124-805-475

Query Match 99.2%; Score 2578; DB 14; Length 1002;
Best Local Similarity 99.6%; Pred. No. 5.8e-140;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALAEIEMQKSVPNKALELNEQTLRADEILPS 60
Db 338 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALAEIEMQKSVPNKALELNEQTLRADEILPS 397

Qy 61 ESKQDYESSWDSLSLCTETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 398 ESKQDYESSWDSLSLCTETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457

Qy 121 KVSIPITKALELMDMOTFKAEPPKPSAFEPALAEIEMQKSVPNKALELNEQTLRADEILPS 180
Db 458 KVSIPITKALELMDMOTFKAEPPKPSAFEPALAEIEMQKSVPNKALELNEQTLRADEILPS 517

Qy 181 SKQKVEENSWDSLSLCTETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 240
Db 518 SKQKVEENSWDSLSLCTETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 577

Qy 241 RARELQKHCEQRTGMEQMKKFCVLLKKLSEAKEIKSLEQENQKVKWEQELCSVRLTLN 300
Db 578 RARELQKHCEQRTGMEQMKKFCVLLKKLSEAKEIKSLEQENQKVKWEQELCSVRLTLN 637

Qy 301 QEEKRRNADILNEKIREELGRIBEQHREKELEVKQLEQALRIQDIELKSVESNLNOVSH 360
Db 638 QEEKRRNADILNEKIREELGRIBEQHREKELEVKQLEQALRIQDIELKSVESNLNOVSH 697

Qy 361 THENENYLLHENCMLKKEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 698 THENENYLLHENCMLKKEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 757

Qy 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHQI 480
Db 758 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHQI 817

Qy 481 VTSRKSQEPAFHAGDACLQKMNVDVSSST 510
Db 818 VTSRKSQEPAFHAGDACLQKMNVDVSSST 847

Search completed: February 19, 2005, 02:33:41
Job time : 50.0894 secs

CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 310, 429, 522
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-076-622-475

Query Match 99.2%; Score 2578; DB 14; Length 1002;
Best Local Similarity 99.6%; Pred. No. 5.8e-140;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALAEIEMQKSVPNKALELNEQTLRADEILPS 60
Db 338 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALAEIEMQKSVPNKALELNEQTLRADEILPS 397

Qy 61 ESKQDYESSWDSLSLCTETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 398 ESKQDYESSWDSLSLCTETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457

Qy 121 KVSIPITKALELMDMOTFKAEPPKPSAFEPALAEIEMQKSVPNKALELNEQTLRADEILPS 180
Db 458 KVSIPITKALELMDMOTFKAEPPKPSAFEPALAEIEMQKSVPNKALELNEQTLRADEILPS 517

Qy 181 SKQKVEENSWDSLSLCTETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 240
Db 518 SKQKVEENSWDSLSLCTETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 577

Qy 241 RARELQKHCEQRTGMEQMKKFCVLLKKLSEAKEIKSLEQENQKVKWEQELCSVRLTLN 300
Db 578 RARELQKHCEQRTGMEQMKKFCVLLKKLSEAKEIKSLEQENQKVKWEQELCSVRLTLN 637

Qy 301 QEEKRRNADILNEKIREELGRIBEQHREKELEVKQLEQALRIQDIELKSVESNLNOVSH 360
Db 638 QEEKRRNADILNEKIREELGRIBEQHREKELEVKQLEQALRIQDIELKSVESNLNOVSH 697

Qy 361 THENENYLLHENCMLKKEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 698 THENENYLLHENCMLKKEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 757

Qy 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHQI 480
Db 758 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHQI 817

Qy 481 VTSRKSQEPAFHAGDACLQKMNVDVSSST 510
Db 818 VTSRKSQEPAFHAGDACLQKMNVDVSSST 847

RESULT 15
US-10-124-805-475
; Sequence 475, Application US/10124805
; Publication No. US20030166022A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C12
; CURRENT APPLICATION NUMBER: US/10/124,805
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapiens

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:52:33 ; Search time 55.7765 Seconds
(without alignments)
3550.262 Million cell updates/sec

Title: US-09-602-362E-16

Perfect score: 2599

Sequence: 1 MKVSIPTKALEMDMQTFKA.....IAGDACLQRKNVDVSTDI 512

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2599	100.0	512	4 AAB84701	AAB84701 Amino aci
2	2589	99.6	1341	4 AAB84702	AAB84702 Amino aci
3	2589	99.6	1341	5 ABJ05537	ABJ05537 Breast ca
4	2589	99.6	1341	6 ABJ37784	ABJ37784 Human tum
5	2589	99.6	1341	6 ABJ47548	ABJ47548 Breast ca
6	2589	99.6	1341	7 ADL93227	ADL93227 Human bre
7	2589	99.6	1349	6 ABJ37788	ABJ37788 Human tum
8	2589	99.6	1349	7 ADL93235	ADL93235 Human bre
9	2578	99.2	1002	4 AAU33351	AAU33351 Human bre
10	2578	99.2	1002	5 ABG78918	ABG78918 Human bre
11	2578	99.2	1002	6 ABJ37741	ABJ37741 Human tum
12	2578	99.2	1002	7 ADL93137	ADL93137 Human bre
13	2578	99.2	1002	8 ADE44427	ADE44427 Human bre
14	2578	99.2	1095	4 AAU33357	AAU33357 Human bre
15	2578	99.2	1095	5 ABG78924	ABG78924 Human bre
16	2578	99.2	1095	6 ABJ37747	ABJ37747 Human tum
17	2578	99.2	1095	7 ADL93155	ADL93155 Human bre
18	2576	99.1	1013	6 ABJ37783	ABJ37783 Human tum
19	2576	99.1	1013	7 ADL93215	ADL93215 Human bre
20	2125	81.8	445	4 AAB50249	AAB50249 Human bre
21	2125	81.8	445	4 AAG65987	AAG65987 B726P spl
22	2125	81.8	445	4 AAU33350	AAU33350 Human bre
23	2125	81.8	445	5 ABG78917	ABG78917 Human bre
24	2125	81.8	445	6 ABJ37740	ABJ37740 Human tum
25	2125	81.8	445	7 ADL93135	ADL93135 Human bre

ALIGNMENTS

RESULT 1

AAB84701
ID AAB84701 standard; protein; 512 AA.

XX AC AAB84701;

XX DT 17-SEP-2001 (first entry)

XX DE Amino acid sequence of a human cancer associated antigen.

XX KW Cancer associated antigen; INGI; tumour suppressor; cancer; vaccine.

XX OS Homo sapiens.

XX PN WO200147959-A2.

XX PD 05-JUL-2001.

XX PF 29-NOV-2000; 2000WO-US042334.

XX PR 30-NOV-1999; 99US-00451739.

XX PR 24-OCT-2000; 2000US-00602362.

XX PA (LUDW-) LUDWIG INST CANCER RES.
PA (SLOK) SLOAN KETTERING INST CANCER RES.
XX (CORR) CORNELL RES FOUND INC.

XX PI Jager D, Stockert E, Scanlan M, Knuth A, Old L, Gure A, Chen Y;

XX WPI; 2001-441706/47.

XX N-PSDB; AAH28489.

XX DR Isolated cancer associated nucleic acid molecule identified by SEREX
DR (serological identification of antigens by recombinant expression
PT cloning) technique, useful in nucleic acid based therapies to treat
PT cancer.

XX PS Claim 83; Page 50-51; 62pp; English.

XX CC The present sequence represents a human cancer associated antigen. The
CC sequence was identified using probes derived from the INGI gene. The INGI
CC gene is a tumour suppressor candidate gene. The cancer associated antigen
CC polynucleotides and polypeptides are useful for screening for the
CC possible presence of a pathological condition in a subject such as
CC cancer. The cancer associated antigen polypeptides are useful for
CC producing vaccines

XX SQ Sequence 512 AA;

Ade44425 Human bre
Ade47547 Breast ca
Abj37789 Human tum
Adl93239 Human bre
Aab84703 Amino aci
Adj69751 Human hea
Aag65983 B726P spl
Aau33346 Human bre
Abg78913 Human bre
Abj37736 Human tum
Adl93131 Human bre
Ade44421 Human bre
Aau33358 Human bre
Abg78925 Human bre
Abj37748 Human tum
Adl93156 Human bre
Abj37782 Human tum
Adl93214 Human bre
Aab07638 Amino aci

Query Match 100.0%; Score 2599; DB 4; Length 512;
Best Local Similarity 100.0%; Pred. No. 1.9e-172;
Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSIPTKALELMDMOTFKAEPPKPSAFEPKSAPEPAIEMQKSVNPKALELNEQTLRADELPLS 60
DB 1 MKVSIPTKALELMDMOTFKAEPPKPSAFEPKSAPEPAIEMQKSVNPKALELNEQTLRADELPLS 60

QY 61 ESKQDYESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
DB 61 ESKQDYESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120

QY 121 KVSIPTKALELMDMOTFKAEPPKPSAFEPKSAPEPAIEMQKSVNPKALELNEQTLRADEMPSE 180
DB 121 KVSIPTKALELMDMOTFKAEPPKPSAFEPKSAPEPAIEMQKSVNPKALELNEQTLRADEMPSE 180

QY 181 SKQKVEENSWDSLSLRETTSQKDVCPKATHQKEMDKISGKLEDSTSLSKILDTVHSCE 240
DB 181 SKQKVEENSWDSLSLRETTSQKDVCPKATHQKEMDKISGKLEDSTSLSKILDTVHSCE 240

QY 241 RARELQDHCEQRTGKMEQMKKFCVLKXKLSEAKEIKSQLENQKVWEQELCSVRLTLN 300
DB 241 RARELQDHCEQRTGKMEQMKKFCVLKXKLSEAKEIKSQLENQKVWEQELCSVRLTLN 300

QY 301 QEEKRRNADILNEKIREELGRIBEEQHRKELEVQKQLEQALRIQDIELKSVEENLNQVSH 360
DB 301 QEEKRRNADILNEKIREELGRIBEEQHRKELEVQKQLEQALRIQDIELKSVEENLNQVSH 360

QY 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
DB 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420

QY 421 KEESLTGRASQYSGQLKVLIAENTMTLSKLEKQDKLEIAEIESHHPRLASAVQDHDQI 480
DB 421 KEESLTGRASQYSGQLKVLIAENTMTLSKLEKQDKLEIAEIESHHPRLASAVQDHDQI 480

QY 481 VTSRKSQEPAFHIAAGDACLQKMNVDVSSTDI 512
DB 481 VTSRKSQEPAFHIAAGDACLQKMNVDVSSTDI 512

RESULT 2
AAB84702
ID AAB84702 standard; protein; 1341 AA.
XX AC
XX AC
XX AAB84702;
XX 17-SEP-2001 (first entry)
XX DE Amino acid sequence of a human cancer associated antigen.
XX KW Cancer associated antigen; INGI; tumour suppressor; cancer; vaccine.
XX OS Homo sapiens.
XX FN WO200147959-A2.
XX PD 05-JUL-2001.
XX PF 29-NOV-2000; 2000WO-US042334.
XX PR 30-NOV-1999; 99US-00451739.
XX PR 24-OCT-2000; 2000US-00602362.
XX XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX PA (SLOK) SLOAN KETTERING INST CANCER RES.
XX PA (CORR) CORNELL RES FOUND INC.
XX FI Jager D, Stockert E, Scanlan M, Knuth A, Old L, Gure A, Chen Y;
XX DR WPI; 2001-441706/47.
XX XX

PT Isolated cancer associated nucleic acid molecule identified by SEREX
PT (serological identification of antigens by recombinant expression
PT cloning) technique, useful in nucleic acid based therapies to treat
PT cancer.
XX
XX
PS Claim 83; Page 53-57; 62pp; English.
XX
XX The present sequence represents a human cancer associated antigen. The
CC sequence was identified using probes derived from the INGI gene. The INGI
CC gene is a tumour suppressor candidate gene. The cancer associated antigen
CC polynucleotides and polypeptides are useful for screening for the
CC possible presence of a pathological condition in a subject such as
CC cancer. The cancer associated antigen polypeptides are useful for
CC producing vaccines
XX
SQ Sequence 1341 AA;

Query Match 99.6%; Score 2589; DB 4; Length 1341;
Best Local Similarity 100.0%; Pred. No. 3.3e-171;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSIPTKALELMDMOTFKAEPPKPSAFEPKSAPEPAIEMQKSVNPKALELNEQTLRADELPLS 60
DB 677 MKVSIPTKALELMDMOTFKAEPPKPSAFEPKSAPEPAIEMQKSVNPKALELNEQTLRADELPLS 736

QY 61 ESKQDYESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
DB 737 ESKQDYESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 796

QY 121 KVSIPTKALELMDMOTFKAEPPKPSAFEPKSAPEPAIEMQKSVNPKALELNEQTLRADEMPSE 180
DB 797 KVSIPTKALELMDMOTFKAEPPKPSAFEPKSAPEPAIEMQKSVNPKALELNEQTLRADEMPSE 856

QY 181 SKQKVEENSWDSLSLRETTSQKDVCPKATHQKEMDKISGKLEDSTSLSKILDTVHSCE 240
DB 857 SKQKVEENSWDSLSLRETTSQKDVCPKATHQKEMDKISGKLEDSTSLSKILDTVHSCE 916

QY 241 RARELQDHCEQRTGKMEQMKKFCVLKXKLSEAKEIKSQLENQKVWEQELCSVRLTLN 300
DB 917 RARELQDHCEQRTGKMEQMKKFCVLKXKLSEAKEIKSQLENQKVWEQELCSVRLTLN 976

QY 301 QEEKRRNADILNEKIREELGRIBEEQHRKELEVQKQLEQALRIQDIELKSVEENLNQVSH 360
DB 977 QEEKRRNADILNEKIREELGRIBEEQHRKELEVQKQLEQALRIQDIELKSVEENLNQVSH 1036

QY 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
DB 1037 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 1096

QY 421 KEESLTGRASQYSGQLKVLIAENTMTLSKLEKQDKLEIAEIESHHPRLASAVQDHDQI 480
DB 1097 KEESLTGRASQYSGQLKVLIAENTMTLSKLEKQDKLEIAEIESHHPRLASAVQDHDQI 1156

QY 481 VTSRKSQEPAFHIAAGDACLQKMNVDVSST 510
DB 1157 VTSRKSQEPAFHIAAGDACLQKMNVDVSST 1186

RESULT 3
AABJ05537
ID AABJ05537 standard; protein; 1341 AA.
XX AC
XX AABJ05537;
XX 14-NOV-2002 (first entry)
XX DE Breast cancer-associated protein 2.
XX DE Breast cancer; breast cancer-associated gene sequence; drug development;
XX KW pharmacogenetics; biosensor development.
XX OS Unidentified.
XX XX

Db 737 ESKQDYESSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLESPDNDGFLKAPCRM 796
Qy 121 KVSIPTKALELMDMOTFKAEPEKPSAFEPALIEKQSVNPKALELKNQTLRADQMPPE 180
Db 797 KVSIPTKALELMDMOTFKAEPEKPSAFEPALIEKQSVNPKALELKNQTLRADQMPPE 856
Qy 181 SKQKVEENSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLESPDNDGFLKAPCRM 240
Db 857 SKQKVEENSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLESPDNDGFLKAPCRM 916
Qy 241 RARELQDHCCEQRTGKMEQMKKFCVLKSLSEAKETKSLQENQKVKWEQELCSVRLTLN 300
Db 917 RARELQDHCCEQRTGKMEQMKKFCVLKSLSEAKETKSLQENQKVKWEQELCSVRLTLN 976
Qy 301 QEEKRRNADILNEKIREELGRIBEOHREKLEVKQLEQALRIQDIELKSVESNLNOVSH 360
Db 977 QEEKRRNADILNEKIREELGRIBEOHREKLEVKQLEQALRIQDIELKSVESNLNOVSH 1036
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 1037 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 1096
Qy 421 KEESLTKRASQYSGQLKVLIAENTMTLSKLEKQDKETLEAEIESHHPRLASAVQDHDQI 480
Db 1097 KEESLTKRASQYSGQLKVLIAENTMTLSKLEKQDKETLEAEIESHHPRLASAVQDHDQI 1156
Qy 481 VTSRKSQEPAPHIAGDACLQKMNVDVSSST 510
Db 1157 VTSRKSQEPAPHIAGDACLQKMNVDVSSST 1186

RESULT 5
ABR47548
ID ABR47548 standard; protein; 1341 AA.
XX ABR47548;
AC ABR47548;
XX 12-JUN-2003 (first entry)
XX Breast cancer associated protein sequence SEQ ID NO:334.
XX Human; breast cancer; cytostatic; gene therapy.
XX Homo sapiens.
XX WO2003004989-A2.
XX 16-JAN-2003.
XX 21-JUN-2002; 2002WO-US019669.
XX 21-JUN-2001; 2001US-0299887P.
XX 27-JUN-2001; 2001US-0301572P.
XX 18-JUL-2001; 2001US-0306501P.
XX 23-SEP-2001; 2001US-0325002P.
XX 05-MAR-2002; 2002US-0362885P.
XX 14-MAY-2002; 2002US-0380391P.
XX (MILL-) MILLENIUM PHARM INC.
XX Lillie J, Gannavarapu M, Glatt K, Hoersch S, Kamatkar S;
XX Mertens M, Monahan JE, Myer Y, Xu Y, Zhao X, Meyers RE;
XX Bast RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;
XX WPI; 2003-210381/20.
XX N-PSDB; ACC50246.
XX Breast cancer diagnosis or treatment by comparing the level of expression
XX of a marker in a patient sample with that in the control non-breast
XX cancer sample.
XX Claim 1; SEQ ID NO 334; 128pp; English.

XX The present invention describes a method for assessing whether a patient
CC is afflicted with breast cancer. The method comprises comparing the level
CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
CC ABR47386 to ABR47632) in a patient sample and the normal level of
CC expression of the marker in a control non-breast cancer sample, where a
CC significant increase in the level of expression of the marker in the
CC patient sample and the normal level is an indication that the patient is
CC afflicted with breast cancer. The breast cancer associated sequences from
CC the present invention have cytostatic activities and can be used in gene
CC therapy. The method is useful for diagnosing and treating breast cancer.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1341 AA;

Query Match 99.6%; Score 2589; DB 6; Length 1341;
Best Local Similarity 100.0%; Pred. No. 3.3e-171;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMOTFKAEPEKPSAFEPALIEKQSVNPKALELKNQTLRADQMPPE 60
Db 677 MKVSIPTKALELMDMOTFKAEPEKPSAFEPALIEKQSVNPKALELKNQTLRADQMPPE 736
Qy 61 ESKQDYESSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLESPDNDGFLKAPCRM 120
Db 737 ESKQDYESSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLESPDNDGFLKAPCRM 796
Qy 121 KVSIPTKALELMDMOTFKAEPEKPSAFEPALIEKQSVNPKALELKNQTLRADQMPPE 180
Db 797 KVSIPTKALELMDMOTFKAEPEKPSAFEPALIEKQSVNPKALELKNQTLRADQMPPE 856
Qy 181 SKQKVEENSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLESPDNDGFLKAPCRM 240
Db 857 SKQKVEENSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLESPDNDGFLKAPCRM 916
Qy 241 RARELQDHCCEQRTGKMEQMKKFCVLKSLSEAKETKSLQENQKVKWEQELCSVRLTLN 300
Db 917 RARELQDHCCEQRTGKMEQMKKFCVLKSLSEAKETKSLQENQKVKWEQELCSVRLTLN 976
Qy 301 QEEKRRNADILNEKIREELGRIBEOHREKLEVKQLEQALRIQDIELKSVESNLNOVSH 360
Db 977 QEEKRRNADILNEKIREELGRIBEOHREKLEVKQLEQALRIQDIELKSVESNLNOVSH 1036
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 1037 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 1096
Qy 421 KEESLTKRASQYSGQLKVLIAENTMTLSKLEKQDKETLEAEIESHHPRLASAVQDHDQI 480
Db 1097 KEESLTKRASQYSGQLKVLIAENTMTLSKLEKQDKETLEAEIESHHPRLASAVQDHDQI 1156
Qy 481 VTSRKSQEPAPHIAGDACLQKMNVDVSSST 510
Db 1157 VTSRKSQEPAPHIAGDACLQKMNVDVSSST 1186

RESULT 6

ADL93227
ID ADL93227 standard; protein; 1341 AA.
XX ADL93227;
AC ADL93227;
XX 20-MAY-2004 (first entry)
DT 20-MAY-2004 (first entry)
XX Human breast cancer-associated polypeptide #49.
XX gene therapy; protein therapy; vaccine; breast cancer; cancer; human.
XX Homo sapiens.
XX US2003166022-A1.

XX PD 04-SEP-2003.
XX XX 15-APR-2002; 2002US-00124805.
XX PF 28-DEC-1998; 98US-00222575.
XX PR 02-APR-1999; 99US-00285480.
XX PR 23-JUN-1999; 99US-00339338.
XX PR 02-SEP-1999; 99US-00389681.
XX PR 03-NOV-1999; 99US-00433826.
XX PR 17-APR-2000; 2000US-00551621.
XX PR 08-JUN-2000; 2000US-00590751.
XX PR 22-JUN-2000; 2000US-00604287.
XX PR 20-JUL-2000; 2000US-00620405.
XX PR 13-APR-2001; 2001US-00834759.
XX PR 07-DEC-2001; 2001US-00007805.
XX PR 13-FEB-2002; 2002US-00076622.
XX FA (CORI-) CORIXA CORP.
XX XX Houghton RL, Sleath PR, Persing DH;
XX XX WPI: 2003-874918/81.
XX DR N-PSDB; ADL93226.
XX XX
XX PT An isolated oncogenic polypeptide useful for preventing, diagnosing and
XX PT treating breast cancer.
XX XX
XX PS Example 12; SEQ ID NO 565; 294pp; English.
XX XX
XX CC The invention relates to an isolated breast cancer-associated
XX CC polypeptide. The polypeptide may be used for the diagnosis and treatment
XX CC of breast cancers. The methods are useful for detecting the presence of a
XX CC cancer in a patient and treating a cancer in a patient. The present
XX CC sequence represents the amino acid sequence of a human breast cancer-
XX CC associated polypeptide.
XX XX
XX SQ Sequence 1341 AA;
Query Match 99.6%; Score 2589; DB 7; Length 1341;
Best Local Similarity 100.0%; Pred. No. 3.3e-171; Indels 0; Gaps 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMTFKAEPPKPSAFEPALMOKSVPNKALELKNQETLRADILPS 60
Db 677 MKVSIPTKALELMDMTFKAEPPKPSAFEPALMOKSVPNKALELKNQETLRADILPS 736
Qy 61 ESKQDYESSWDSSELCETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 737 ESKQDYESSWDSSELCETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 796
Qy 121 KVSIPTKALELMDMTFKAEPPKPSAFEPALMOKSVPNKALELKNQETLRADMPFSE 180
Db 797 KVSIPTKALELMDMTFKAEPPKPSAFEPALMOKSVPNKALELKNQETLRADMPFSE 856
Qy 181 SKQKVEENSWDSSELRSTVSQKDVCPKATHQKEMDKISGKLESDTSLSKILTVDHSC 240
Db 857 SKQKVEENSWDSSELRSTVSQKDVCPKATHQKEMDKISGKLESDTSLSKILTVDHSC 916
Qy 241 RARELQKDHCHQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQLCSVRLTLN 300
Db 917 RARELQKDHCHQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQLCSVRLTLN 976
Qy 301 QEEKRRNADLTNKKIREEELGRIEEOHREKLEVKQOQALRIODIELKSVESNLNOVSH 360
Db 977 QEEKRRNADLTNKKIREEELGRIEEOHREKLEVKQOQALRIODIELKSVESNLNOVSH 1036
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 1037 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 1096
Qy 421 KEESLTKRASQYSGQLKVLIANTWLTSLKKEQDKTEILEAIESHHPRLASAVQDHQOI 480

Db 1097 KEESLTKRASQYSGQLKVLIANTWLTSLKKEQDKTEILEAIESHHPRLASAVQDHQOI 1156
Qy 481 VTSRKQEPAFHIAAGDAQLQRKMNVDVSS 510
Db 1157 VTSRKQEPAFHIAAGDAQLQRKMNVDVSS 1186
RESULT 7
ABJ37788
ID ABJ37788 standard; protein; 1349 AA.
XX AC ABJ37788;
XX XX
XX DT 15-MAY-2003 (first entry)
XX XX Human tumour-related protein - SEQ ID No 573.
XX DE Human; vaccine; gene therapy; T cell stimulation; T cell expansion;
XX KW tumour; breast cancer; cancer; immune response stimulation.
XX XX Homo sapiens.
XX OS WO200283956-A1.
XX PN 24-OCT-2002.
XX PD 15-APR-2002; 2002WO-US012378.
XX PF 13-APR-2001; 2001US-00834759.
XX PR 07-DEC-2001; 2001US-00007805.
XX PR 13-FEB-2002; 2002US-00076622.
XX XX (CORI-) CORIXA CORP.
XX XX Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
XX PI Mitcham JR, Xu J, Harlocker SL, Hepler WI, Henderson RA, Fanger GR;
XX PI Vedvick TS, McNeill PD, Durham M;
XX DR WPI: 2003-103376/09.
XX XX
XX PT New polypeptide and polynucleotide useful for stimulating and/or
XX PT expanding T cells specific for a tumor protein and treating breast
XX PT cancer.
XX PS Example 12; Page 353-357; 375pp; English.
XX XX
XX CC The invention comprises a method of stimulating and/or expanding T cells
XX CC specific for a tumour protein. The invention further comprises human
XX CC nuclear acids and proteins that are associated with tumours (e.g. breast
XX CC cancer). The method and sequences of the invention are useful for
XX CC stimulating and/or expanding T cells specific for a tumour protein,
XX CC detecting the presence of cancer, stimulating an immune response in a
XX CC patient and treating breast cancer. The present amino acid sequence
XX CC represents a human tumour-related protein
XX XX
XX SQ Sequence 1349 AA;
Query Match 99.6%; Score 2589; DB 6; Length 1349;
Best Local Similarity 100.0%; Pred. No. 3.4e-171; Indels 0; Gaps 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMTFKAEPPKPSAFEPALMOKSVPNKALELKNQETLRADILPS 60
Db 685 MKVSIPTKALELMDMTFKAEPPKPSAFEPALMOKSVPNKALELKNQETLRADILPS 744
Qy 61 ESKQDYESSWDSSELCETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 745 ESKQDYESSWDSSELCETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 804
Qy 121 KVSIPTKALELMDMTFKAEPPKPSAFEPALMOKSVPNKALELKNQETLRADMPFSE 180
Db 805 KVSIPTKALELMDMTFKAEPPKPSAFEPALMOKSVPNKALELKNQETLRADMPFSE 864

QY	181	SKQKVEENSWDSRLRTVSQKDVCPKATHQKEMDKISGKLEDSLSKILDTVHSCE	240
DB	865	SKQKVEENSWDSRLRTVSQKDVCPKATHQKEMDKISGKLEDSLSKILDTVHSCE	924
QY	241	RARELQKHCEORTGKQMKKFCVLKKLSEAKEIKSOLENKVKWEQELCSVRLTLN	300
DB	925	RARELQKHCEORTGKQMKKFCVLKKLSEAKEIKSOLENKVKWEQELCSVRLTLN	984
QY	301	QBEKKRNADILNEKIREELGRIBEQHREKELEVKQQLAEQALRIQDIELKSVESNLQVSH	360
DB	985	QBEKKRNADILNEKIREELGRIBEQHREKELEVKQQLAEQALRIQDIELKSVESNLQVSH	1044
QY	361	THENENYLHENCMLKKEIAMLKLEIATLKHOYEKENKYPEDIKILKEKNAELQMTLKL	420
DB	1045	THENENYLHENCMLKKEIAMLKLEIATLKHOYEKENKYPEDIKILKEKNAELQMTLKL	1104
QY	421	KEESLTKRASQSGQLKVLIAENTMLTSKLEKQDKKEILEAEIESHHPRLASAVODHDOI	480
DB	1105	KEESLTKRASQSGQLKVLIAENTMLTSKLEKQDKKEILEAEIESHHPRLASAVODHDOI	1164
QY	481	VTSRKQEPAFHAGDACLQKKNVDVVSST	510
DB	1165	VTSRKQEPAFHAGDACLQKKNVDVVSST	1194
RESULT 8			
ADL93235	ID	ADL93235 standard; protein; 1349 AA.	
AC	AC	ADL93235;	
XX	AC	ADL93235;	
XX	DT	20-MAY-2004 (first entry)	
XX	DE	Human breast cancer-associated polypeptide #53.	
XX	KW	gene therapy; protein therapy; vaccine; breast cancer; cancer; human.	
XX	OS	Homo sapiens.	
XX	PN	US2003166022-A1.	
XX	PD	04-SEP-2003.	
XX	PF	15-APR-2002; 2002US-00124805.	
XX	PR	28-DEC-1998; 98US-00222575.	
XX	PR	02-APR-1999; 95US-00285480.	
XX	PR	23-JUN-1999; 99US-00339338.	
XX	PR	02-SEP-1999; 99US-00389681.	
XX	PR	03-NOV-1999; 99US-00433826.	
XX	PR	17-APR-2000; 2000US-00551621.	
XX	PR	08-JUN-2000; 2000US-00590751.	
XX	PR	22-JUN-2000; 2000US-00604287.	
XX	PR	20-JUL-2000; 2000US-00620405.	
XX	PR	13-APR-2001; 2001US-00834759.	
XX	PR	07-DEC-2001; 2001US-00007805.	
XX	PR	13-FEB-2002; 2002US-00076622.	
XX	PA	(CORI-) CORIXA CORP.	
XX	PI	Houghton RL, Sleath PR, Persing DH;	
XX	DR	WPI; 2003-874918/81.	
XX	DR	N-PSDB; ADL93228.	
XX	PT	An isolated oncogenic polypeptide useful for preventing, diagnosing and treating breast cancer.	
XX	PS	Example 12; SEQ ID NO 573; 294pp; English.	
XX	XX	The invention relates to an isolated breast cancer-associated polypeptide. The polypeptide may be used for the diagnosis and treatment of a patient with breast cancer, the presence of a	

PA (CORI-) CORIXA CORP.
XX Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
XX WPI: 2001-611721/70.
XX N-PSDB; AAS47411.
PT Breast Tumor Proteins and nucleic acids useful for the prevention,
XX diagnosis and treatment of breast cancer.
XX Claim 3; Page 281-283; 297pp; English.
XX The invention relates to isolated breast tumour proteins and nucleic
XX acids that encode them, including immunogenic fragments of the proteins.
XX Also included are expression vectors expressing the proteins, transformed
XX cells and antibodies raised against the proteins or an antigen presenting
XX cell expressing the protein. The proteins and nucleic acids may be used
XX in the prevention, diagnosis and treatment of diseases associated with
XX inappropriate breast tumour protein expression, i.e. breast tumours and
XX breast cancer e.g by gene therapy. The nucleic acids and their
XX complements may also be used as DNA probes in diagnostic assays to detect
XX and quantitate the presence of similar nucleic acids in samples, and
XX therefore which patients may be in need of restorative therapy. The
XX proteins, nucleic acids and antibodies may be used in assays to identify
XX modulators (e.g. antagonists) of breast tumour protein expression and
XX activity. The antibodies and antagonists may also be used to down
XX regulate expression and activity. The antibodies may also be used as
XX diagnostic agents for detecting the presence of the proteins in samples
XX (e.g. by enzyme linked immunosorbant assay (ELISA)) and in other immuno-
XX purification diagnostic techniques. The present sequence is a breast
XX tumour protein encoded by a cDNA from a breast tumour cDNA library
XX isolated by subtractive hybridisation against a normal breast cDNA
XX library
SQ Sequence 1002 AA;

Query Match 99.2%; Score 2578; DB 4; Length 1002;
Best Local Similarity 99.6%; Pred. No. 1.3e-170;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFAPAEIEMQKSVNPKALELKNQETLRADIELPS 60
Db 338 MKVSIPTKALELMDMOTFKAEPPKPSAFAPAEIEMQKSVNPKALELKNQETLRADIELPS 397
Qy 61 ESKQKDYBESSWDSLSLCTETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 398 ESKQKDYBESSWDSLSLCTETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457
Qy 121 KVSIPTKALELMDMOTFKAEPPKPSAFAPAEIEMQKSVNPKALELKNQETLRADQMPFSE 180
Db 458 KVSIPTKALELMDMOTFKAEPPKPSAFAPAEIEMQKSVNPKALELKNQETLRADQMPFSE 517
Qy 181 SKQKXVEENSWSLSRLTTSQKDVCPKATHQKEMDKISGLBEDSTSLKILDTVHSC 240
Db 518 SKQKXVEENSWSLSRLTTSQKDVCPKATHQKEMDKISGLBEDSTSLKILDTVHSC 577
Qy 241 RARELQKDHCQRTGKMKQMKKFCVLKKLSEAKEIKSQLENQKVKWEQLCSVRLTLN 300
Db 578 RARELQKDHCQRTGKMKQMKKFCVLKKLSEAKEIKSQLENQKVKWEQLCSVRLTLN 637
Qy 301 QEEKRRNADILNEKIREELGRIBEQHRLKEVQKQLEALRIQDIELKSVESNLNQVSH 360
Db 638 QEEKRRNADILNEKIREELGRIBEQHRLKEVQKQLEALRIQDIELKSVESNLNQVSH 697
Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENYFEDIKILKEKNAELQMTLKL 420
Db 698 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENYFEDIKILKEKNAELQMTLKL 757
Qy 421 KEESLTKRASQVSGQLKVLIAENTMLTSLKLEKQDKKEILEAIESHHPRLASAVQDHDQI 480
Db 758 KEESLTKRASQVSGQLKVLIAENTMLTSLKLEKQDKKEILEAIESHHPRLASAVQDHDQI 817
Qy 481 VTSRKSQBPAPHIAGDAQLQRKMNVDSST 510

Db 818 VTSRKSQBPAPHIAGDAQLQRKMNVDSST 847
RESULT 10
ABG78918
ID ABG78918 standard; protein; 1002 AA.
XX AC ABG78918;
XX DT 15-NOV-2002 (first entry)
XX DE Human breast tumour polypeptide #10.
XX KW Human; breast tumour protein; breast cancer; cytostatic; vaccine.
XX OS Homo sapiens.
XX PN US2002085998-A1.
XX PD 04-JUL-2002.
XX PF 13-APR-2001; 2001US-00834759.
XX PR 28-DEC-1998; 98US-00222575.
XX PR 02-APR-1999; 99US-00285480.
XX PR 23-JUN-1999; 99US-00339338.
XX PR 02-SEP-1999; 99US-00389681.
XX PR 03-NOV-1999; 99US-00433826.
XX PR 17-APR-2000; 2000US-00551621.
XX PR 08-JUN-2000; 2000US-00590751.
XX PR 22-JUN-2000; 2000US-00604287.
XX PR 20-JUL-2000; 2000US-00620405.
XX PA (CORI-) CORIXA CORP.
XX PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
XX FI Henderson RA;
XX WPI: 2002-635657/68.
XX N-PSDB; ABS64012.
PT Novel breast cancer polynucleotides and polypeptides encoded by the
XX polynucleotides, useful for detecting the presence of breast cancer in a
XX patient, and in pharmaceutical compositions, for treating breast cancer.
XX Claim 2; Page 208-211; 247pp; English.
XX The invention relates to an isolated breast tumour polynucleotide and the
XX polypeptide it encodes. The polynucleotide and polypeptide are useful for
XX detecting the presence of breast cancer in a patient, and in
XX pharmaceutical compositions for treating breast cancer. The sequences are
XX useful for stimulating an immune response in a patient and can therefore
XX be used in production of vaccines. The sequences are also useful for
XX detecting the presence of a cancer in a patient, by obtaining a
XX biological sample from the patient, contacting the biological sample with
XX a composition of the invention and detecting the amount of polynucleotide
XX that hybridizes to the sample. This sequence represents a human breast
XX tumour polypeptide of the invention
SQ Sequence 1002 AA;

Query Match 99.2%; Score 2578; DB 5; Length 1002;
Best Local Similarity 99.6%; Pred. No. 1.3e-170;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFAPAEIEMQKSVNPKALELKNQETLRADIELPS 60
Db 338 MKVSIPTKALELMDMOTFKAEPPKPSAFAPAEIEMQKSVNPKALELKNQETLRADIELPS 397
Qy 61 ESKQKDYBESSWDSLSLCTETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 398 ESKQKDYBESSWDSLSLCTETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457

PR 17-APR-2000; 2000US-00551621.
PR 08-JUN-2000; 2000US-00590751.
PR 22-JUN-2000; 2000US-00604287.
PR 20-JUL-2000; 2000US-00620405.
PR 13-APR-2001; 2001US-00834759.
PR 07-DEC-2001; 2001US-00007805.
PR 13-FEB-2002; 2002US-00076622.
XX (CORI-) CORIXA CORP.
XX Houghton RL, Sleath PR, Persing DH;
XX WPI; 2003-874918/81.
XX N-PSDB; ADL93136.
XX An isolated oncogenic polypeptide useful for preventing, diagnosing and
XX treating breast cancer.
XX Example 1; SEQ ID NO 475; 294pp; English.
XX The invention relates to an isolated breast cancer-associated
XX polypeptide. The polypeptide may be used for the diagnosis and treatment
XX of breast cancers. The methods are useful for detecting the presence of a
XX cancer in a patient and treating a cancer in a patient. The present
XX sequence represents the amino acid sequence of a human breast cancer-
XX associated polypeptide.
XX Sequence 1002 AA;
Query Match 99.2%; Score 2578; DB 7; Length 1002;
Best Local Similarity 99.6%; Pred. No. 1.3e-170;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMTFKAEPPKPSAFEPKPAIEQKSVNPKALELKNQETLRADIELPS 60
Db 338 MKVSIPTKALELMDMTFKAEPPKPSAFEPKPAIEQKSVNPKALELKNQETLRADIELPS 397
Qy 61 ESKQDYEESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 398 ESKQDYEESSWDSLSLCTVTSQKDVCLPKAXHQKEIDKINGKLEESPDNDGFLKAPCRM 457
Qy 121 KVSIPTKALELMDMTFKAEPPKPSAFEPKPAIEQKSVNPKALELKNQETLRADMFPS 180
Db 458 KVSIPTKALELMDMTFKAEPPKPSAFEPKPAIEQKSVNPKALELKNQETLRADMFPS 517
Qy 181 SKQKVEENSWDSLSLRETQSKDVCPKATHQKEMDKISGLSDSTLSKILDTVHSC 240
Db 518 SKQKVEENSWDSLSLRETQSKDVCPKATHQKEMDKISGLSDSTLSKILDTVHSC 577
Qy 241 RARELQKDHCHQRTGRMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 300
Db 578 RARELQKDHCHQRTGRMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 637
Qy 301 QEEKRRNADTLNKEIRBELGRIBEQHKELEVKQLEQALRIQDIELKSVEENLNQVSH 360
Db 638 QEEKRRNADTLNKEIRBELGRIBEQHKELEVKQLEQALRIQDIELKSVEENLNQVSH 697
Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLHQYQEKENKYFEDIKILKEKNAELQMTLKL 420
Db 698 THENENYLLHNCMLKKEIAMLKLEIATLHQYQEKENKYFEDIKILKEKNAELQMTLKL 757
Qy 421 KEESLTKRASQYSQLKVLIAENTMLTSKLKEKQDKLEIAEIESHPRLASAVQDDHQI 480
Db 758 KEESLTKRASQYSQLKVLIAENTMLTSKLKEKQDKLEIAEIESHPRLASAVQDDHQI 817
Qy 481 VTSRKSQEPHFIAGDAQLQRKMNVDVSS 510
Db 818 VTSRKSQEPHFIAGDAQLQRKMNVDVSS 847

RESULT 13
ADE44427
ID ADE44427 standard; protein; 1002 AA.

XX ADE44427;
XX 29-JAN-2004 (first entry)
XX Human breast cancer protein #10.
XX human; breast tumour; cancer; vaccine; T cell stimulator;
XX T cell expander.
XX Homo sapiens.
XX US2003104366-A1.
XX 05-JUN-2003.
XX 17-APR-2000; 2000US-00551621.
XX 28-DEC-1998; 98US-00222575.
XX 02-APR-1999; 99US-00285480.
XX 23-JUN-1999; 99US-00339338.
XX 02-SEP-1999; 99US-00389681.
XX 03-NOV-1999; 99US-00433826.
XX (JIAN/) JIANG Y.
XX (DILL/) DILLON D C.
XX (MITC/) MITCHAM J L.
XX (XUJJ/) XU J.
XX (HARL/) HARLOCKER S L.
XX Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL;
XX WPI; 2004-020270/02.
XX N-PSDB; ADE44426.
XX Novel isolated polypeptide comprising immunogenic portion of breast tumor
XX protein or its variant, useful for formulating vaccines for inhibiting
XX cancer development in a patient.
XX Example 1; SEQ ID NO 475; 217pp; English.
XX The invention relates to an isolated polypeptide comprising at least an
XX immunogenic portion of a breast tumour protein. The polynucleotide, its
XX polypeptide, its antibody, a pharmaceutical composition comprising the
XX fusion protein or the polynucleotide encoding it, a vaccine comprising
XX the fusion protein or the polynucleotide encoding it, an isolated T cell
XX population comprising T cells specific for a breast tumour protein, and a
XX method for removing tumour cells from a biological sample is useful for
XX inhibiting the development of a cancer in a patient. The polypeptide is
XX useful for stimulating and/or expanding T cells specific for a breast
XX tumour protein. Stimulating and/or expanding T cells specific for a
XX breast tumour protein is useful for inhibiting the development of a
XX cancer in a patient. The method additionally involves the step of cloning
XX at least one proliferated cell and then administering the cloned T cells
XX to the patient. The present sequence represents a breast cancer protein.
XX Sequence 1002 AA;
Query Match 99.2%; Score 2578; DB 8; Length 1002;
Best Local Similarity 99.6%; Pred. No. 1.3e-170;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMTFKAEPPKPSAFEPKPAIEQKSVNPKALELKNQETLRADIELPS 60
Db 338 MKVSIPTKALELMDMTFKAEPPKPSAFEPKPAIEQKSVNPKALELKNQETLRADIELPS 397
Qy 61 ESKQDYEESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 398 ESKQDYEESSWDSLSLCTVTSQKDVCLPKAXHQKEIDKINGKLEESPDNDGFLKAPCRM 457
Qy 121 KVSIPTKALELMDMTFKAEPPKPSAFEPKPAIEQKSVNPKALELKNQETLRADMFPS 180
Db 458 KVSIPTKALELMDMTFKAEPPKPSAFEPKPAIEQKSVNPKALELKNQETLRADMFPS 517

CC and quantitate the presence of similar nucleic acids in samples, and
CC therefore which patients may be in need of restorative therapy. The
CC proteins, nucleic acids and antibodies may be used in assays to identify
CC modulators (e.g. antagonists) of breast tumour protein expression and
CC activity. The antibodies and antagonists may also be used to down
CC regulate expression and activity. The antibodies may also be used as
CC diagnostic agents for detecting the presence of the proteins in samples
CC (e.g. by enzyme linked immunosorbant assay (ELISA)) and in other immuno-
CC purification diagnostic techniques. The present sequence is a breast
CC tumour protein encoded by a cDNA from a breast tumour cDNA library
CC isolated by subtractive hybridisation against a normal breast cDNA
CC library
CC
SQ Sequence 1095 AA;
Query Match 99.2%; Score 2578; DB 4; Length 1095;
Best Local Similarity 99.6%; Pred. NO. 1.5e-170; Indels 0; Gaps 0;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MKVSIPTKALELMDMOTFKAEPEKPSAFEPALTEMQKSVNPKALELKNQTLRADPILPS 60
Db 431 MKVSIPTKALELMDMOTFKAEPEKPSAFEPALTEMQKSVNPKALELKNQTLRADPILPS 490
QY 61 ESKQDYESSWDSSESLCETVSKQDVCLPKAKHQEKIDKINGKLESPNDGFLKAPCRM 120
Db 491 ESKQDYESSWDSSESLCETVSKQDVCLPKAKHQEKIDKINGKLESPNDGFLKAPCRM 550
QY 121 KVSIPPTKALELMDMOTFKAEPEKPSAFEPALTEMQKSVNPKALELKNQTLRADPILPS 180
Db 551 KVSIPPTKALELMDMOTFKAEPEKPSAFEPALTEMQKSVNPKALELKNQTLRADPILPS 610
QY 181 SKQKVEENSWSSESLRETYSQKDVCPKATHQKEMDKISGKLESDSTLSKILDTVHSCE 240
Db 611 SKQKVEENSWSSESLRETYSQKDVCPKATHQKEMDKISGKLESDSTLSKILDTVHSCE 670
QY 241 RARELQKHCEORTGKMEQMKKFCVLKKLSEAKETKSOLENQKVKWEQELCSVRLTLN 300
Db 671 RARELQKHCEORTGKMEQMKKFCVLKKLSEAKETKSOLENQKVKWEQELCSVRLTLN 730
QY 301 QEEKRRNADILNEKIREELGRIEEQRKELEVKQQLQEQALRIQDIELKSVESNLNQVSH 360
Db 731 QEEKRRNADILNEKIREELGRIEEQRKELEVKQQLQEQALRIQDIELKSVESNLNQVSH 790
QY 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKL 420
Db 791 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKL 850
QY 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHDQI 480
Db 851 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHDQI 910
QY 481 VTSRKSQEPAFHAGDACLQKMNVDVSST 510
Db 911 VTSRKSQEPAFHAGDACLQKMNVDVSST 940
RESULT 15
ABG78924
ID ABG78924 standard; protein; 1095 AA.
XX AC ABG78924;
XX DT 15-NOV-2002 (first entry)
XX DE Human breast tumour polypeptide #15.
XX KW Human; breast tumour protein; breast cancer; cytostatic; vaccine.
XX OS Homo sapiens.
XX PN US2002085998-A1.
XX PD 04-JUL-2002.
QY 181 SKQKVEENSWSSESLRETYSQKDVCPKATHQKEMDKISGKLESDSTLSKILDTVHSCE 240
Db 518 SKQKVEENSWSSESLRETYSQKDVCPKATHQKEMDKISGKLESDSTLSKILDTVHSCE 577
QY 241 RARELQKHCEORTGKMEQMKKFCVLKKLSEAKETKSOLENQKVKWEQELCSVRLTLN 300
Db 578 RARELQKHCEORTGKMEQMKKFCVLKKLSEAKETKSOLENQKVKWEQELCSVRLTLN 637
QY 301 QEEKRRNADILNEKIREELGRIEEQRKELEVKQQLQEQALRIQDIELKSVESNLNQVSH 360
Db 638 QEEKRRNADILNEKIREELGRIEEQRKELEVKQQLQEQALRIQDIELKSVESNLNQVSH 697
QY 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKL 420
Db 698 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKL 757
QY 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHDQI 480
Db 758 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHDQI 817
QY 481 VTSRKSQEPAFHAGDACLQKMNVDVSST 510
Db 818 VTSRKSQEPAFHAGDACLQKMNVDVSST 847
RESULT 14
AAU33357
ID AAU33357 standard; protein; 1095 AA.
XX AC AAU33357;
XX DT 18-DEC-2001 (first entry)
XX DE Human breast cancer protein B726P fusion protein #1.
XX KW Human; ss; breast cancer protein; tumour; cancer; cytostatic;
XX KW gene therapy; immunogen.
XX OS Homo sapiens.
XX PN WO200179286-A2.
XX PD 25-OCT-2001.
XX PF 12-APR-2001; 2001WO-US012164.
XX PR 17-APR-2000; 2000US-00551621.
XX PR 08-JUN-2000; 2000US-00590751.
XX PR 22-JUN-2000; 2000US-00604287.
XX PR 20-JUL-2000; 2000US-00620405.
XX PA (CORI-) CORIXA CORP.
XX PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
XX WPI; 2001-611721/70.
XX DR N-PSDB; AAS47421.
XX PT Breast Tumor Proteins and nucleic acids useful for the prevention,
XX PT diagnosis and treatment of breast cancer.
XX PS Claim 22; Page 292-295; 297pp; English.
XX CC The invention relates to isolated breast tumour proteins and nucleic
XX CC acids that encode them, including immunogenic fragments of the proteins.
XX CC Also included are expression vectors expressing the proteins, transformed
XX CC cells and antibodies raised against the proteins or an antigen presenting
XX CC cell expressing the protein. The proteins and nucleic acids may be used
XX CC in the prevention, diagnosis and treatment of diseases associated with
XX CC inappropriate breast tumour protein expression, i.e. breast tumours and
XX CC breast cancer e.g. by gene therapy. The nucleic acids and their
XX CC complements may also be used as DNA probes in diagnostic assays to detect

Qy 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLEKQDKKEILEAEIESHHPRLASAVQDHQOI 480
Db 851 KEESLTKRASQYSGQLKVLIAENTMLTSKLEKQDKKEILEAEIESHHPRLASAVQDHQOI 910
Qy 481 VTSRKSQEPAPHIAGDACLOKQKMNVDVSST 510
Db 911 VTSRKSQEPAPHIAGDACLOKQKMNVDVSST 940

Search completed: February 19, 2005, 02:04:38
Job time : 58.7765 secs

XX 13-APR-2001; 2001US-00834759.
XX 28-DEC-1998; 98US-00222575.
PR 02-APR-1999; 99US-00285480.
PR 23-JUN-1999; 99US-00339338.
PR 02-SEP-1999; 99US-00389661.
PR 03-NOV-1999; 99US-00433826.
PR 17-APR-2000; 2000US-00551621.
PR 08-JUN-2000; 2000US-00590751.
PR 22-JUN-2000; 2000US-00604287.
PR 20-JUL-2000; 2000US-00620405.
XX (CORI-) CORIXA CORP.
XX Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
PI Henderson RA;
XX WPI; 2002-635657/68.
DR N-PSDB; ABS64022.
XX Novel breast cancer polynucleotides and polypeptides encoded by the
PT polynucleotides, useful for detecting the presence of breast cancer in a
PT patient, and in pharmaceutical compositions, for treating breast cancer.
XX Disclosure; Page 223-225; 247pp; English.
XX The invention relates to an isolated breast tumour polynucleotide and the
CC polypeptide it encodes. The polynucleotide and polypeptide are useful for
CC detecting the presence of breast cancer in a patient, and in
CC pharmaceutical compositions for treating breast cancer. The sequences are
CC useful for stimulating an immune response in a patient and can therefore
CC be used in production of vaccines. The sequences are also useful for
CC detecting the presence of a cancer in a patient, by obtaining a
CC biological sample from the patient, contacting the biological sample with
CC a composition of the invention and detecting the amount of polynucleotide
CC that hybridizes to the sample. This sequence represents a human breast
CC tumour polypeptide of the invention
XX Sequence 1095 AA;
SQ
Query Match 99.2%; Score 2578; DB 5; Length 1095;
Best Local Similarity 99.6%; Pred. No. 1.5e-170;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMOTFKAEPPPEKPSAFEPALIEQKSVNPKALELKNQTLRADEILPS 60
Db 431 MKVSIPTKALELMDMOTFKAEPPPEKPSAFEPALIEQKSVNPKALELKNQTLRADEILPS 490
Qy 61 ESKQDYESSWDSSESLCETVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 491 ESKQDYESSWDSSESLCETVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 550
Qy 121 KVSIPTKALELMDMOTFKAEPPPEKPSAFEPALIEQKSVNPKALELKNQTLRADOMPFSE 180
Db 551 KVSIPTKALELMDMOTFKAEPPPEKPSAFEPALIEQKSVNPKALELKNQTLRADOMPFSE 610
Qy 181 SKQKVENSWDSSESLRETTSQKDVCPKATHQKEMDKISGKLESDTSLSKILDTVHSCE 240
Db 611 SKQKVENSWDSSESLRETTSQKDVCPKATHQKEMDKISGKLESDTSLSKILDTVHSCE 670
Qy 241 RARELQKDHCQRTGKMEQMKKFCVLKKLSSEAKEIKSQLENQKVKWEQELCSVRLTLN 300
Db 671 RARELQKDHCQRTGKMEQMKKFCVLKKLSSEAKEIKSQLENQKVKWEQELCSVRLTLN 730
Qy 301 QEEKRRNADILNEKIRELGRIBEQHREKELEVKQQLQALRIQDIELKSVESNINQVSH 360
Db 731 QEEKRRNADILNEKIRELGRIBEQHREKELEVKQQLQALRIQDIELKSVESNINQVSH 790
Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHOYEKENKYPEDIKILKEKNAELQMTLKL 420
Db 791 THENENYLLHNCMLKKEIAMLKLEIATLKHOYEKENKYPEDIKILKEKNAELQMTLKL 850

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:55:34 ; Search time 124.257 Seconds
(without alignments)
4166.462 Million cell updates/sec

Title: US-09-602-362E-27
Perfect score: 5173
Sequence: 1 MVATLLSYGAVIEYQNKASL.....REQLKVLTAENTMLTSLKE 1011

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_eprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5173	100.0	1011	2 Q9BXX2	Q9bxx2 homo sapien
2	3075	59.4	1341	2 Q9BXX3	Q9bxx3 homo sapien
3	1011.5	19.6	424	2 Q9NS19	Q9ns19 homo sapien
4	923	17.8	1715	2 Q9UPS8	Q9ups8 homo sapien
5	920.5	17.8	1710	2 Q9HIQ1	Q9hiq1 homo sapien
6	703	13.6	1080	2 Q9ZRI4	Q9zri4 homo sapien
7	553	10.7	168	2 Q9N7B8	Q9n7b8 homo sapien
8	489	9.5	718	2 Q9HCD1	Q9hcd1 homo sapien
9	465	9.0	1043	2 Q9ZS22	Q9zs22 mus musculu
10	396	7.7	823	2 Q9H0H6	Q9h0h6 homo sapien
11	359	6.9	2006	2 Q7K5Q6	Q7k5q6 plasmodium
12	359	6.9	2019	2 Q7K5Q5	Q7k5q5 plasmodium
13	359	6.9	2055	2 Q8IHP3	Q8ihp3 plasmodium
14	359	6.9	2055	2 Q8TSC7	Q8tsc7 plasmodium
15	357.5	6.9	10578	2 Q8ISF5	Q8isf5 caenorhabdi
16	357.5	6.9	18519	2 Q8ISP6	Q8isf6 caenorhabdi
17	357.5	6.9	18534	2 Q8ISP7	Q8isf7 caenorhabdi
18	346.5	6.7	2042	2 Q8BUQ9	Q8buq9 debaryomyce
19	346	6.7	1416	2 Q9BZF9	Q9bzf9 homo sapien
20	346	6.7	1449	2 Q9BG87	Q9bg87 bos taurus
21	345	6.7	1416	2 Q9HCL1	Q9hcl1 homo sapien
22	342.5	6.6	1980	2 Q6FWE0	Q6fwe0 candida gla
23	341.5	6.6	1427	1 R8ST_HUMAN	P30622 homo sapien
24	340.5	6.6	1268	2 Q7LGT1	Q7lgt1 saccharomyc
25	340.5	6.6	1790	2 Q07380	Q07380 saccharomyc
26	340	6.6	1413	2 Q8CGB3	Q8cgb3 mus musculu
27	338.5	6.5	1790	1 U501_YEAST	P25386 saccharomyc
28	337	6.5	641	2 Q60311	Q60311 homo sapien
29	336	6.5	453	2 Q8IZM7	Q8izm7 homo sapien
30	336	6.5	992	1 AN18_HUMAN	Q8ivf6 homo sapien
31	331.5	6.4	7210	2 Q9V7G8	Q9v7g8 drosophila

RESULT 1

Q9BXX2	PRELIMINARY;	PRT;	1011	AA.
AC	Q9BXX2			
DT	01-JUN-2001 (Tremblrel. 17, Created)			
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)			
DT	01-JUN-2003 (Tremblrel. 24, Last annotation update)			
DE	Breast cancer antigen NY-BR-1.1 (Fragment).			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21174979; PubMed=11280766;			
RA	Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karbach J.,			
RA	Jaeger E., Knuth A., Old L.J., Chen Y.T.;			
RT	"Identification of a Tissue-specific Putative Transcription Factor in			
RT	Breast Tissue by Serological Screening of a Breast Cancer Library.;"			
RL	Cancer Res. 61:2055-2061(2001)			
DR	EMBL; AF269088; AAK27326.1; -			
DR	HSSP; O75832; IUOH.			
DR	InterPro; IPR002110; ANK.			
DR	Pfam; PF00023; Ank; 3.			
DR	SMART; SM00248; ANK; 3.			
DR	PROSITE; PS50088; ANK_REPEAT; 1.			
DR	PROSITE; PS50297; ANK_REPEAT_REGION; 1.			
KW	ANK repeat.			
FT	NON TER 1011			
SQ	SEQUENCE 1011 AA; 114250 MW; 16627D7B218D8438 CRC64;			
Query Match 100.0%; Score 5173; DB 2; Length 1011;				
Best Local Similarity 100.0%; Pred. No. 1.6e-194;				
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	MVATLLSYGAVIEYQNKASLTPLLLAIQKRSKQTVFELLTKNANANAFNESKCTALMLAI	60	
Db	1	MVATLLSYGAVIEYQNKASLTPLLLAIQKRSKQTVFELLTKNANANAFNESKCTALMLAI	60	
Qy	61	CEGSSEIVGMLLQNVDFVPAEDIHGITAERYAAARGVNYIHQQLLEHIRKLKPKPQNTNP	120	
Db	61	CEGSSEIVGMLLQNVDFVPAEDIHGITAERYAAARGVNYIHQQLLEHIRKLKPKPQNTNP	120	
Qy	121	EGTGTGTPDEAPLAERTPDTABSLLEKTPDEARLVEGTSKAKIQCLGKATSGKFEQSTE	180	
Db	121	EGTGTGTPDEAPLAERTPDTABSLLEKTPDEARLVEGTSKAKIQCLGKATSGKFEQSTE	180	
Qy	181	ETPRKILRPTKETSEKFSWPAKERSRKITWEEKTSVKTECVAGVTPNKTVELEKGTSM	240	
Db	181	ETPRKILRPTKETSEKFSWPAKERSRKITWEEKTSVKTECVAGVTPNKTVELEKGTSM	240	
Qy	241	IACPTKETSTKASTNVDVSSVEPIFSLFGTRTINSQCTKVEEDFNLATKIISKSAQNY	300	
Db	241	IACPTKETSTKASTNVDVSSVEPIFSLFGTRTINSQCTKVEEDFNLATKIISKSAQNY	300	

ALIGNMENTS

Qy	301	TCLPDATYQKDIKTINHKIEDQMPSPSKREDEDEYSDWDSGLSPSSAKTOVCIPESMYQ	360
Db	301	TCLPDATYQKDIKTINHKIEDQMPSPSKREDEDEYSDWDSGLSPSSAKTOVCIPESMYQ	360
Qy	361	KVMEINREVELPEKPSAFKPAVEMQKTPVNPNAFELKNEQTLRAAQMPFSPESKQKDDEN	420
Db	361	KVMEINREVELPEKPSAFKPAVEMQKTPVNPNAFELKNEQTLRAAQMPFSPESKQKDDEN	420
Qy	421	SWDSSEPCETVQSKDVYLPKATHQKEFTLISGKLEESPVKDGLLKPTCGRKVSLPNKALE	480
Db	421	SWDSSEPCETVQSKDVYLPKATHQKEFTLISGKLEESPVKDGLLKPTCGRKVSLPNKALE	480
Qy	481	LKDRETFKAEPSDKDGLLKPTCGRKVSLPNKALEKDBETLKASPDNDGLLKPTCGRKV	540
Db	481	LKDRETFKAEPSDKDGLLKPTCGRKVSLPNKALEKDBETLKASPDNDGLLKPTCGRKV	540
Qy	541	SLPNKALELKDRETFKAAQMPSPSKOKDDENSWDFSEFLETLQNDVCLPKATHQKEF	600
Db	541	SLPNKALELKDRETFKAAQMPSPSKOKDDENSWDFSEFLETLQNDVCLPKATHQKEF	600
Qy	601	DTLSGKLEESPDKGLLKPTCGMKISLPNKALELKDRETFKAEVSSVESTFSFGKPTT	660
Db	601	DTLSGKLEESPDKGLLKPTCGMKISLPNKALELKDRETFKAEVSSVESTFSFGKPTT	660
Qy	661	ENSQTKVEEDPNLTKEGATKTVTGOQERDGIIEIRAPQDQTNKMPSELGRKEDTKST	720
Db	661	ENSQTKVEEDPNLTKEGATKTVTGOQERDGIIEIRAPQDQTNKMPSELGRKEDTKST	720
Qy	721	SDSEIISVSDTONYECPLPEATYQKEIKITNGKIEBSPKSHFEPATEMNSVFNKGLEW	780
Db	721	SDSEIISVSDTONYECPLPEATYQKEIKITNGKIEBSPKSHFEPATEMNSVFNKGLEW	780
Qy	781	KVKQTLRADSTTSLKILDALPSCERGBRELKDNCEQITAKMEQMKVFCVLQKELSEAKE	840
Db	781	KVKQTLRADSTTSLKILDALPSCERGBRELKDNCEQITAKMEQMKVFCVLQKELSEAKE	840
Qy	841	IKSOLENQAKWEQELCSVRLPLNQBEEKRRNVDLKEKIRPEOLRKKLVKQLEQTL	900
Db	841	IKSOLENQAKWEQELCSVRLPLNQBEEKRRNVDLKEKIRPEOLRKKLVKQLEQTL	900
Qy	901	RQDIELKSVTSNLNQVSHSENDLFHENCMLKEIAMLKLEVALTLKHQHVKENYKF	960
Db	901	RQDIELKSVTSNLNQVSHSENDLFHENCMLKEIAMLKLEVALTLKHQHVKENYKF	960
Qy	961	EDIKILQENKAELOMTLKLKQKTVTKRASQYREQLKVLTAENTMLTSKLKE	1011
Db	961	EDIKILQENKAELOMTLKLKQKTVTKRASQYREQLKVLTAENTMLTSKLKE	1011

RESULT 2

Q9BXX3

PRELIMINARY; PRT; 1341 AA.

Q9BXX3

AC Q9BXX3;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Breast cancer antigen NY-BR-1.

DE Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN (1)

RP SEQUENCE FROM N.A.

RX MEDLINE=21174979; PubMed=11280766;

RA Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karbach J.,

RA Jaeger E., Knuth A., Old L.J., Chen Y.T.;

RT "Identification of a Tissue-specific Putative Transcription Factor in

RT Breast Tissue by Serological Screening of a Breast Cancer Library.;"

RL Cancer Res. 61:2035-2061(2001).

DR ENBL; AF269087; AAK27325.1; --

DR GO; GO:0005634; C:nucleus; NAS.

DR GO; GO:0005515; F:protein binding; NAS.

DR	GO; GO:0003700; F:transcription factor activity; NAS.		
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.		
DR	InterPro; IPR002110; ANK.		
DR	InterPro; IPR001969; Pept_Asp_AS.		
DR	Pfam; PF00023; Ank; 6.		
DR	PRINTS; PR01415; ANKYRIN.		
DR	SMART; SM00248; ANK; 6.		
DR	PROSITE; PS00088; ANK_REPEAT; 4.		
DR	PROSITE; PS0297; ANK_REPEAT_REGION; 1.		
DR	PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.		
KW	ANK repeat.		
SQ	SEQUENCE 1341 AA; 152776 MW; 33ES3DDBE6FD3A58B CRC64;		
Query Match	59.4%; Score 3075; DB 2; Length 1341;		
Best Local Similarity	57.4%; Pred. No. 2.3e-112;		
Matches	664; Conservative 76; Mismatches 145; Indels 272; Gaps 12;		
QY	1	MVATLSYGAVIEVQNKASITPLLLAIQKRSKQTVFLLTKNANAFVSKCTALMLAI	60
DB	98	VWAKLLSHGAVIEVHNKASITPLLLSITKRSEQIVFLLTKNANAVNKYKCTALMLAV	157
QY	61	CEGSSEIVGMLLOQNVDFAEIDHIGITAEYAAAGVNYIHQOLLEHTRKLPKNPONTNP	120
DB	158	CHGSSEIVGMLLOQNVDFPAADICGVTAHYAVTCGFHHIHEQIMYIRKLSKHQNTNP	217
QY	121	EGTGTGTPDEAAPLA-----ERTPTDAESLLEKTPDEAARLVE	158
DB	218	EGTSAGTPDEAAPLAERTPTDAESLVEKTPDEAAPLVERPTDAESLVEKTPDEAASLVE	277
QY	159	GTSAKIQCLKATSGKFEQSTETPRKILAPPTKETSEKFSWP-----	200
DB	278	GTSDKIQCLEKATSGKFEQSAEETPREITSPAKETSEKFTWPAKGRPKTAWKEDTTPR	337
QY	201	-----AKERSKITWEEKETSVKTECVAGVTNKTVELEKGTSNMIACP	244
DB	338	EIMSPAKETSEKFTWAAKGRPKLANEKETPVKTGCVARVTSNKTIVLEKGRSKMIACP	397
QY	245	TKETSTKASTNDVSSVEPIPSLFGTRTNIENSQCTKVEEDFNLATKIISKAAQNYTCLP	304
DB	398	TKESSTKASAN-----	408
QY	305	DATYQKDIKTINHKIEDQMPFSPSKREDEDEYSDWDSGLSPFESSAKTOVCIPESMYQKVE	364
DB	409	-----DQRFPSKQEBDEYSCDSRLSPFESSAKIOVCIPESIQKVE	452
QY	365	INREVELPEKPSAFKPAVEMQKTPVNPNAFELKNEQTLRAAQMPFSPESKQKDDENSWDS	424
DB	453	INREVELPEKPSAFKPAIEMQNSVPNKAFLKNEQTLRADPMFPFSPESKQKDYENSWDS	512
QY	425	ESPCETVQSKDVYLPKATHQKEFTLISGKLEESPVKDGLLKPTCGRKVSLPNKALELADR	484
DB	513	ESLCEVTSQKDVCLPKATHQKEIDKINGKLEESPNKDGLLKATCGMKVSIPTKALELADM	572
QY	485	ETFKAESPDKGLLKPTCGRKVSLPNKALELKDBETLKAESPDNDGLLKPTCGRKVSLPN	544
DB	573	QTFKAEPSPKPSAFEP-----ATMQK-----SVPN	598
QY	545	KALELKDBETFKAAQMPFSPESKQKDDENSWDFLETLQNDVCLPKATHQKEFTLS	604
DB	599	KALELKNEQTLRADEILLPSPESKQKDYENSWDFSELCETVQSKDVCLPKAAHQKEIDKN	658
QY	605	GKLEESPDKGLLKPTCGMKISLPNKALELKDBETFKAEVSSVESTFSFGKPTTNSQ	664
DB	659	GKLEGSVPKDGLLKANCCKMVSIPTKALELMDMTFKAAPPEK-PSAF-----EPAIEMQK	713
QY	665	STKVEEDFNLTKEGATKTVTGOQERDGIIEIRAPQDQTNKMPSELGRKEDTKSTDSSE	724
DB	714	SV-----PNKALELKNEQTLRADEILL-----SESKQKDYESSWDSSE	751
QY	725	IISVSDTONYECPLPEATYQKEIKITNGKIEB-----	756
DB	752	SLCEVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMQ	811


```

Qy 757 -----PEKSPHFPATQMSPVKNGLKWKQKTLRA----- 788
Db 812 TFKAEPPEKPSAFPAIEMQKSVENKALELNEQTLRADQMPPSESQKKVVEENSWDS 871
Qy 789 -----DSTTLKSLDALPSCERGLKNDKCEQITA 819
Db 872 LRETVSQDVCVPKATQKEMDKISGLKEDSTLSKILDTVHSCERARELQKHCQRTG 931
Qy 820 KMEQMKKFCVLOKELSEAKEIKSOLENQAKWQELCSVRLPLNQBEERKRNVDILKEK 879
Db 932 KMEQMKKFCVLLKKLSEAKEIKSOLENQKVKWQELCSVRLTLNQBEERKERNADILNEK 991
Qy 880 IRP-----EEQLRKLEVKHOLEQTLRIQDIELKSVTSNLSNQHSHSENDLFHENCML 934
Db 992 IREBLGRIEEQRKELEVKQLEQALRIQDIELKSVESNLSNQHSHSENNYLLHENCML 1051
Qy 935 KKETAMKLEVLTKHOFKOVENKVFEDIKLOKQNAELQMTLKKOKTVTKRASQYREQ 994
Db 1052 KKEIAMKLEIATUKHQVEKENKVFEDIKILKRNALQMTLKKESLTKRASQYGGQ 1111
Qy 995 LKVLTAENTMLTSKLKE 1011
Db 1112 LKVLIAENTMLTSKLKE 1128

RESULT 3
Q9NSI9 ID Q9NSI9 PRELIMINARY; PRT; 424 AA.
AC Q9NSI9;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE PRED4 protein (fragment).
GN Name=PRED4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RZ MEDLINE=99397452; PubMed=10470851;
RA Kikuno R., Nagase T., Ishikawa K., Hirose M., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:197-205 (1999).
DR EMBL; AB028997; BAA83026.2; -.
DR HSP; P20749; IKIB.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ANK; 4.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS00088; ANK_REPEAT; 4.
DR PROSITE; PSS0297; ANK_REPEAT; 1.
KW ANK repeat.
FT NON TER
SQ SEQUENCE 424 AA; 50038 MW; C6E704795C534625 CRC64;

Query Match 19.6%; Score 1011.5; DB 2; Length 424;
Best Local Similarity 91.5%; Pred. No. 2.9e-32;
Matches 204; Conservative 6; Mismatches 34; Indels 9; Gaps 1;

Qy 789 DSTTLKSLDALPSCERGLKNDKCEQITAKMQKNKFCVLOKELSEAKEIKSOLENQ 848
Db 2 DSTTLKSLDALPSCERGLKNDKCEQITAKMQKNKFCVLOKELSEAKEIKSOLENQ 61
Qy 849 KAKWEQELCSVRLPLNQBEERKRNVDILKEKIRPEQRLKKLEVKHOLEQTLRIQDIELK 908
Db 62 KAKWEQELCS-----KKRDVDILKEKIRPEQRLKKLEVKHOLEQALRIQDIELK 112
Qy 909 SVTSNLSNQHSHSENDLFHENCMLKKEIAMKLEVLTKHOFKOVENKVFEDIKILQ 968
Db 113 SVTSNLSNQHSHSENDLFHENCMLKKEIAMKLEVLTKHOFKOVENKVFEDIKILQ 172

```

```

Qy 969 KNAELQMTLKLKOKTVTKRASQYREQKLVLTAEINTMLTSKLKE 1011
Db 173 KNAELQMTLKLKOKTVTKRASQYREQKLVLTAEINTMLTSKLKE 215

RESULT 4
Q9UPS8 ID Q9UPS8 PRELIMINARY; PRT; 1715 AA.
AC Q9UPS8;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE KIAA1074 protein (fragment).
GN Name=KIAA1074;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RZ MEDLINE=99397452; PubMed=10470851;
RA Kikuno R., Nagase T., Ishikawa K., Hirose M., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:197-205 (1999).
DR EMBL; AB028997; BAA83026.2; -.
DR HSP; P20749; IKIB.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ANK; 4.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS00088; ANK_REPEAT; 4.
DR PROSITE; PSS0297; ANK_REPEAT; 1.
KW ANK repeat.
FT NON TER
SQ SEQUENCE 1715 AA; 196962 MW; A828589F5F58E3203 CRC64;

Query Match 17.8%; Score 923; DB 2; Length 1715;
Best Local Similarity 28.5%; Pred. No. 4.2e-28;
Matches 320; Conservative 157; Mismatches 342; Indels 304; Gaps 34;

Qy 7 SYGAVIEVQ-----NKASLTPLLAIQKRSKOTVBELLTKNANANAFESK 53
Db 61 SAGNVKQVQIILLRKNGLNDROKNTALHLACANGHPEVVTLLVDRKCOLNVCDNR 120
Qy 54 TALMLAICEGSEITVGMLLQONVDVFAEDIHGITAERYAAARGVNYIHQQLLEHRLPK 113
Db 121 TALMKAVQCEKCATILSHGADPNLADVHGNLHVA-----VYNEIDISVATKLL 173
Qy 114 NPQNTNPGTSTGTPDEAAPIAERTPTAESLLEKTDEAARLVEGTSIAQLGKATSG 173
Db 174 YDANIEAK-----NKODLTPL-----LL-----AVSG 195
Qy 174 KFGSTEETPRKILRPKTSEKSPWPAKRSRKITWEEKETSVKTECVAGVTNKTVEL 233
Db 196 KQOQVMEFLRK--KANVNAVDKL-----ESSHOLISEYKEER----- 232
Qy 234 EKGTSNNIACPTKETSTKASTNVDSVVEPFIISLFGTRTIENSQCTKVEEDFNLATKIS 293
Db 233 -----PKHSQNSNVDESSEDSLSRLSGKPGVDDSWPTSDDEDLNFDTKNVP 280
Qy 294 K-----SAAQNYTCUPDATYQKDIKTIHKKI--EDQMPFSESREDEEYVSWDSGLPE 345
Db 281 KPSLAKLMTASQOSRKNLEATY-GTVRTGNRTLPED-----RDSQSDQEVVVSLEPT 331
Qy 346 SSAKTQVCIPESMYQ-----KVMENRREVEELPEKPSAFKP--AVEMQKTVNKA 393
Db 332 TSIKQV-CFHSPTYQSPDLLPKPSHKSHPGLMKKEPTKPGIAKKNGLDIIESAP--- 387

```

```
QY 394 FELKNEOTL-----RAAOMFPSESOKODDEENSWSESPECETVSKQDVYLPKATH 443
Db 388 LEQTNNDNLTVYDEVHKNNRDMMSALGLQGEEDIESPWDSSESINFPPQYV-----440
QY 444 QKEFDTLGKLESPVKGDLKPTCGRKVSLPNKALELKORETTPKASSPDKGGLLKPTCG 503
Db 441 ----DPLAGAAD-----GKEKNIGNEQAE-----DVFYIPSC- 468
QY 504 RKVSLPNKALELKORETL-----KAESPNDGLKPTCGRKVSLPNKALELKORETTPKAA 558
Db 469 MSGSNNFMAKLEDTNRVNGMPVAHMSPERYLHLKPTIEMKDSVPNKGAGMKDVQTSKAA 528
QY 559 Q-----MFPSSKOKDEENSWDFESFLETLQNDVCLPKATHQKEFDTLGKLESPDK 613
Db 529 EHDLEVASSEBQEREGSENNQPQVEERKKNRNMMEVSANIH-----DGATDDAEDDDDD 584
QY 614 DGLLK-----PTCGMKISLPNKALELKORETTPKAAEDVSSVESTFS-LFGKPTT 660
Db 585 DGLIKRKSGETDHOQFPKENKEYASGAPALQMKEVKSTEXEKRTSKESVNSPVFGKASL 644
QY 661 ENSQSTKVEEDFNLTKEGATKTVTGQOERDIGIIERAPOQOTNKMPTSELGRKEDTKST 720
Db 645 LTGGLLQVDDSSLS-----EIDEDEGRPTKTSNEKNKV-----KNQIQSM 686
QY 721 SD-SLIIISVSDTONYEC-LPEATYQKEIKTTNGKIEESPEKPSHFEPATEMONSVPNKG 778
Db 687 DDDVDDLTOSSETASEDCPLHSSYK-----NFMLLIQOLGM 722
QY 779 EWNKQTLRADSTTLKILDALPSCGERELKONCEOITAKMOQMKVFCVLQKELSEA 838
Db 723 ECK-----DSVSLKIQDAALSCEERLELKHONCELLTVKIKKMDKVNVLQRELS 775
QY 839 KEIKSOLENQAKWEQELCSVRLPLNQEEKRRNVDILKEKI-----RPEQJRKLEVK 893
Db 776 KEIKSOLQEHQKWEWERELCSLRFSLNQEEKRRNADTLVEKIREQLRRKEEQYRKEVEV 835
QY 894 HOLEOTLRIQDIELKSVTSNLNOV-----917
Db 836 QLELSLQTLMELETLTVKSNLNOVQERNDQAORQLSRQNAQLQDGLITNLHLSQKEIE 895
QY 918 -----SHTHESENDLPHENCMKKEIAMLKLEVATLKHQHVQKENVYFEDIKILQ 967
Db 896 MAQKMNSENSHSEESKDLGSHKMLQBEIAMLRLDITTKNQNQEKKEKCFEDLKIVK 955
QY 968 EKNALQWLTKLQKTVTKRASOYREQLKVLTAENTMLTSLK 1010
Db 956 EKNEDLQKTIKONEETLTQTTISQYNGRSLVLTAEANMLNSKLE 998

RESULT 5
Q9HIQ1
ID Q9HIQ1 PRELIMINARY; PRT; 1710 AA.
AC Q9HIQ1;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE BA145E8.1 (KIAA1074).
GN Name=BA145E8.1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Bird C.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL162272; CAC19649.1; -.
DR HSSP; P20749; 1K1B.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ANK; 4.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS50088; ANK_REPEAT; 4.
```

```
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW ANK repeat.
SQ SEQUENCE 1710 AA; 196409 MW; 01CBF9BADB894872 CRC64;

Query Match 17.8%; Score 920.5; DB 2; Length 1710;
Best local similarity 28.5%; Pred. No. 5.2e-28;
Matches 320; Conservative 158; Mismatches 341; Indels 305; Gaps 35;

QY 7 SYGAVIEV-----NKASLTPLLLAIQKRSKQTVFELLTKNANANAFNESKC 53
Db 55 SAGWAKVQIILLRRKNGLNDKRNRTALHLACANGHEVTVLLVDRKQCLNVCNENR 114
QY 54 TALMLAICEGSSIVGMLLQONVDVFAEDIHGTASYRYAARGVNIHQOLLEHRIKLPK 113
Db 115 TALMKAVOCQEEKCATILLEHGDADPNLADVHGNTALHYA-----VYNEDISVATKLL 167
QY 114 NPQNTPEGTSTGTPDEAALAEKPTDPAESLEKTPDEAARLVEGTSAKIQCIGKATSG 173
Db 168 YDANIEAK-----NKODLTPL-----LL-----AVSG 189
QY 174 KFEQSTEEETPRKILRPTKETSEKFSWPAKRSRKITWEEKETSVKTECVAGVTPNKTENV 233
Db 190 KQOMVEFLIKK--KANVNAVDKL-----ESSHQLISEYKEERI-----226
QY 234 EKGTSNMIACTKETSTKASTNVDSVPEIFSLFGTRTITENSQCCTKVEEDFNLTAKIIS 293
Db 227 -----PKHSSONSNSVDESSEDSLSRLSGKPGVDSDWPTSDDBDLNFDTKNVP 274
QY 294 K-----SAAQNYTCLPATYQKDIKTINHKI-EQOMFPSESKREDEEYSDSGSLPE 345
Db 275 KPSLAKLMTASQOSRKNLEATY-GTVRTGNRTLTFED-----RSDSQDEVVVESLPT 325
QY 346 SSAKTQVCIPESVMQ-----KWEINRREVEELPEKPSAFKP--AVEMOKTVPNKA 393
Db 326 TSIKVO-CFSPHYQSPDLLPKPSHSLANPLGMKEPTKPGAKKENGDIIESAP---381
QY 394 FELKNEOTL-----RAAOMFPSESOKODDEENSWSESPECETVSKQDVYLPKATH 443
Db 382 LEQTNNDNLTVYDEVHKNNRDMMSALGLQGEEDIESPWDSSESINFPPQYV-----434
QY 444 QKEFDTLGKLESPVKGDLKPTCGRKVSLPNKALELKORETTPKASSPDKGGLLKPTCG 503
Db 435 ----DPLAGAAD-----GKEKNIGNEQAE-----DVFYIPSC- 462
QY 504 RKVSLPNKALELKORETL-----KAESPNDGLKPTCGRKVSLPNKALELKORETTPKAA 558
Db 463 MSGSNNFMAKLEDTNRVNGMPVAHMSPERYLHLKPTIEMKDSVPNKGAGMKDVQTSKAA 522
QY 559 Q-----MFPSSKOKDEENSWDFESFLETLQNDVCLPKATHQKEFDTLGKLESPDK 613
Db 523 EHDLEVASSEBQEREGSENNQPQVEERKKNRNMMEVSANIH-----DGATDDAEDDDDD 578
QY 614 DGLLKPTCGMKI-----SLPNK-----ALELKORETTPKAAEDVSSVESTFS-LFGKPT 659
Db 579 DGLIKRKSGETDHOQFPKENKEYASGAPALQMKEVKSTEXEKRTSKESVNSPVFGKAS 638
QY 660 TENSQSTKVEEDFNLTKEGATKTVTGQOERDIGIIERAPOQOTNKMPTSELGRKEDTKS 719
Db 639 LLTGGLLQVDDSSLS-----EIDEDEGRPTKTSNEKNKV-----KNQIQS 680
QY 720 TSD-SLIIISVSDTONYEC-LPEATYQKEIKTTNGKIEESPEKPSHFEPATEMONSVPNKG 777
Db 681 MDDVDDLTQSETASEDCPLHSSYK-----NFMLLIQOLGM 716
QY 778 LEWNKQTLRADSTTLKILDALPSCGERELKONCEOITAKMOQMKVFCVLQKELSEA 837
Db 717 MECK-----DSVSLKIQDAALSCEERLELKHONCELLTVKIKKMDKVNVLQRELS 769
QY 838 AKETKSOLENQAKWEQELCSVRLPLNQEEKRRNVDILKEKI-----RPEQJRKLEVK 892
Db 770 TKEIKSOLQEHQKWEWERELCSLRFSLNQEEKRRNADTLVEKIREQLRRKEEQYRKEVEV 829
QY 893 KHOLEOTLRIQDIELKSVTSNLNOV-----917
```

```

Db 830 KQOLESQTLMEURTVKSNLNQVQBRNDAQRLSREQNARMLQDGLTNHLSKQKEI 889
Qy 918 -----SHTHESENDFPHENCMKKEIATMLKLEVATLKHOHQVKENKYPEDIKIL 966
Db 890 EMAQCKMNSNSHSEKDLSHKNSMLQOEIAMLRLIEDITKQNOKEKKEKCFEDLKIV 949
Qy 967 QEKNAELQMTLKLKQKTVTKRASQYREQLKVLTAENTMLTSKJK 1010
Db 950 KEKNEDLQTKIQNEETLTQTISQYNGRLSVLTAENAMLSKLE 993

RESULT 6
Q6ZR14 ID Q6ZR14 PRELIMINARY; PRT; 1080 AA.
AC Q6ZR14;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ46736.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Trachea;
RA Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,
RA Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A.,
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK128577; BAC87508.1; -.
FT NON-TER 1080
SQ SEQUENCE 1080 AA; 126741 MW; 47BBA933AE595C CRC64;

Query Match 13.6%; Score 703; DB 2; Length 1080;
Best Local Similarity 32.7%; Pred. No. 9.8e-20;
Matches 197; Conservative 98; Mismatches 163; Indels 144; Gaps 17;

Qy 485 ETFKAESPDKGLKPTCGRKVSLPNKALELKDRETLKAESPDPDGLLKPTCGRKVSLPN 544
Db 17 DAYKTKPIQLNFKQPLDYDHCSSANNYSME-PELENVRSSPRGDT-----SKVSLKE 69
Qy 545 KALELKDBETFK-AAQMPPESSKQKDEENSWDFSEFLETLQNDVCLPKATHQKE----- 599
Db 70 ELQO--DMQRFKNRIGMLKVFEQALEKE-----KVQLQKEVEEERKKHNNMEV 117
Qy 600 ----FDTLGKLESPDKGLKPTCGMKI---SLPNK-----ALELKDBETFK 642
Db 118 SANTHDGATDAEDDDDDGGIIOKRGSETHQFPFKNKEYASSGPAQKEVKSTEK 177
Qy 643 EDVSVSESTFSS-LFGKPTTENSQSTKYVEEDFNLTKEGATKVTGQOERDGIITIERAPOD 701
Db 178 EKRTSKESVNSPVFGKASLLTGGLQLQVDDSSLS-----EIDEDSGRPTKTSN 226
Qy 702 QTNKMPSTSELGRKEDTKSTSD-SHISVSDTQNVEC-LPEATYQKEIKTNGKIESEPEK 759
Db 227 EKNKV-----KNQIQSMDDVDLDTOSSETASEDCELPHSSYK----- 263
Qy 760 PSHPFEPATEMNSVFNKGLKWKNTLRADSTTLTKILDALPSCRGRELKKNCEOITA 819
Db 264 -----NFMLLTEQLGWECK-----DSVSLKLTQDAALSCERLELKNHCELLTV 308
Qy 820 KMEQKMKFVQLKELSEAKEIKSQLENQKAKWEQLCSVRLPLNQBEERKNVDILKEK 879
Db 309 KIKKEDKVVYQLRELSETKEIKSQLEHQKVEWERELCSLRFSLNQBEERKNADTLVEK 368
Qy 980 I-----RPEEQRLKLEVKHOLEOTLRIQDIELKSVTSNMQV----- 917

```

```

Db 369 IREOLRRKEQYRKEVEVKQOLESQTLMEURTVKSNLNQVQBRNDAQRLSREQNA 428
Qy 918 -----SHTHESENDFPHENCMKKEIATMLKLEVATLKHOHQVKENKYPEDIKIL 948
Db 429 RMLQDGLTNHLSKQKEIEMAQCKMNSNSHSEKDLSHKNSMLQOEIAMLRLIEDIT 488
Qy 949 KHQHQVKENKYPEDIKILOEKNAELQMTLKLKQKTVTKRASQYREQLKVLTAENTMLTSK 1008
Db 489 KNQNOKEKKEKCFEDLKIVKEKNEDLQTKIQNEETLTQTISQYNGRLSVLTAENAMLSK 548
Qy 1009 LK 1010
Db 549 LE 550

RESULT 7
Q8N7B8 ID Q8N7B8 PRELIMINARY; PRT; 169 AA.
AC Q8N7B8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ25822.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Musashino K., Yuuki H., Hara H., Suzuki Y., Hata H.,
RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,
RA Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A., Kawakami B.,
RA Nagai K., Isogai T., Sugano S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK098688; BAC05379.1; -.
SQ SEQUENCE 169 AA; 18305 MW; BC6BA50DB58944CF CRC64;

Query Match 10.7%; Score 553; DB 2; Length 168;
Best Local Similarity 73.9%; Pred. No. 8.6e-15;
Matches 113; Conservative 9; Mismatches 27; Indels 4; Gaps 1;

Qy 142 AESLLEKTPDEAARLVVEGTSKIQCLGKATSGKFEQSTETPRKILRPTKTSKFSWPA 201
Db 2 AESLVERPDEAARLVVEGTSKIQCLGKATSGKFEQSAEETPKKIMRTAKETSEKFAWA 61
Qy 202 KERSRKITWEEKETSVKTECVAGVTPNKTVELEKGTSMNIACPTKETSTKASTNVDVSSV 261
Db 62 KERPRKITWEEKETSVKTECVAGVTPNKTVELEKGTSEMLTCTPTKETSTKASTNVDVSSV 121
Qy 262 EPIFSLF-----GTRTIENSQCTKVEEDFNLATK 290
Db 122 ESIFRVSPCHPGWNAMARSRLTATSTSQKEATK 154

RESULT 8
Q9HCD1 ID Q9HCD1 PRELIMINARY; PRT; 718 AA.
AC Q9HCD1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE KIAA1641 protein (Fragment).
GN Name=KIAA1641;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirose M., Ohara O.;

```

RT "Prediction of the coding sequences of unidentified human genes.
 RT XVII. The complete sequences of 100 new cDNA clones from brain which
 RT code for large proteins in vitro."
 RL DNA Res. 7:273-281(2000).
 DR EMBL; AB046861; BAB13467.1; -.
 FT NON TER 1
 SQ SEQUENCE 718 AA; 82972 MW; 2AD05ADE70E3C97F CRC64;

Query Match 9.5%; Score 489; DB 2; Length 718;
 Best Local Similarity 28.5%; Pred. No. 1.4e-11;
 Matches 160; Conservative 79; Mismatches 186; Indels 136; Gaps 12;
 QY 457 SPVKDGLKPTCGRKVSLPNKALELKDETFKASPDGDKGLKPTCGRKVSLPNKALELK 516
 DB 11 SSQKQPAKATSKDKDSVNPTEIKDQCGSGTSSQKQPAWKAITSVKDSVSNIAIEIK 70
 QY 517 DRETLKASPDNDGLKPTCGRKVSLPNKALELKDETFKAAQMPFSESQKQDENSWD 576
 DB 71 DQIRGTVPQKSAQKVFKKVSLNIAIRI-----TGQWK 108
 QY 577 FESFLETLLQNDVCLPKATHQKEFTLSGLKEESPDGDKGLKPTCGRKVSLPNKALELK 636
 DB 109 -----SG--TEYPENLPTLKATIENTKNSVLNATKWK 139
 QY 637 RETFKAEDVSVSTFSLFGKPTTENSQSTKVEDFNLTKEGATKTVTGQOERDIGIIE 696
 DB 140 -----VQTSPTQDLEWASE-----GQKR----- 159
 QY 697 RAPQDTNMPFSELGRKEDTKTSD--SEILSVSTQNYCELPATYQKEIKTTNGKIEE 755
 DB 160 --LEBYENNPQV-----KNQIHSRDDLDIIQSSQTVS----- 191
 QY 756 SPEKSHPEPATQMSVNPNGLEWKNKQTLRADSTTSLKILDALPSCERGELKKNCE 815
 DB 192 -----EDGSLCCNCKNVILLIDQHEMKCKDCVHLKTKNTFCLWKRLIKLDNHCE 243
 QY 816 QITAKMEQWKNKFCVLQKELSEAKIKSQLENQAKWEQELCSVRLPLNQBEERKNVDI 875
 DB 244 QLRVKIRLKNKASVLQKRISEKEIKSQLEKLEKELCSLRFALQEKKKRRNVEE 303
 QY 876 LKEKIR-----PEQRLKLEVKHOLEQTLRIQDIELKSVTSNLNVSHTHESENDLFE 930
 DB 304 LHQVREKLRIITEQYRIEADVTKPIKPAKSAEVELTKTGGNSNQVSETDEKE--DLLHE 362
 QY 931 NCMLKKEIAMLKLEVATLKHQVKNKYFEDIKIQENKAELOMTLKLKQKTVTKRASQ 990
 DB 363 NRLMQDEIARLREKDKTIKNQN--LEKYLKDFEIVKRRKHEDLQALKRNGETLAKTIAC 420
 QY 991 YREQLKVLTAENTMTLSLKE 1011
 DB 421 YSGQLAALTDTENTLRSKLEK 441

RESULT 9

Q992S2 ID Q692S2 PRELIMINARY; PRT; 1043 AA.
 AC Q692S2
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DE MKIAA1074 protein (Fragment).
 GN Name=MKIAA1074;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
 RA Saga Y., Seino S., Nishimura M., Kaisho T., Hoshino K., Kitamura H.,
 RA Nagase T., Ohara O., Koga H.;
 RT "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene:
 IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous

RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
 RT Randomly Sampled from Size-Fractionated Libraries."
 RL DNA Res. 11:205-218(2004).
 DR EMBL; AK173096; BAD32374.1; -.
 DR InterPro; IPR010989; c-snare.
 DR InterPro; IPR009054; Topismrse_insert.
 FT NON TER 1
 SQ SEQUENCE 1043 AA; 121354 MW; 5F37B1F9AA37B212 CRC64;

Query Match 9.0%; Score 465; DB 2; Length 1043;
 Best Local Similarity 33.9%; Pred. No. 1.9e-10;
 Matches 122; Conservative 56; Mismatches 84; Indels 98; Gaps 9;
 QY 701 DQTKMPTSELGRKEDTKTSDSEILSVSTQNYCELPATYQKEIKTTNGKIEESPEKP 760
 DB 91 DQSEAPT-----KTSKKNKDS-----GQTAADVNLDDTFESETASEDELH-QGPDSE 140
 QY 761 SHFEPATEMNSVNPNGLEWKNKQTLRADSTTSLKILDALPSCERGELKKNCEQITAK 820
 DB 141 S-----ILCAIEHLRLCK-----DTASLLKIRDAVSYKRLIELKRSCHCELLTGK 186
 QY 821 MEQWKNKFCVLQKELSEAKIKSQLENQAKWEQELCSVRLPLNQBEERKNVDILKEKI 880
 DB 187 LKRMENKYKGLQKEMSETEEVKSRLEKHEKVGWQELCRLRFALKQEEKKRSADQLSEKT 246
 QY 881 -----RPEQRLKLEVKHOLEQTLRIQDIELKSVTSNLNV----- 917
 DB 247 MEQLRRKGEOQSEVEARQQLSEASLRTLEMLATKVKSHLNQVLEERNETQOLSRQNA 306
 QY 918 -----SH-----THESENDLFHENCMLKKEIAMLKLEVATLK 949
 DB 307 MLQDGLLASHLCKQKEIEMTKQWTSVSVSHEKEKDLKHQRLQDEVAVLREMDTIK 366
 QY 950 HQHQVKNKYFEDIKIQENKAELOMTLKLKQKTVTKRASQYREQLKVLTAENTMTLSKL 1009
 DB 367 SHQKEKRYLEDIKIANEKNDLQRMVKL-----NMLSSKL 403

RESULT 10

Q9H0H6 ID Q9H0H6 PRELIMINARY; PRT; 823 AA.
 AC Q9H0H6
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DE Hypothetical protein DKFZp434A171.
 GN Name=DKFZp434A171;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RG The German cDNA Consortium;
 RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
 RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL136793; CAB66727.1; -.
 DR HSSP; Q60778; IOY3.
 DR InterPro; IPR002110; ANK.
 DR Pfam; PF00023; Ank; 6.
 DR PRINTS; PR01415; ANKYRIN.
 DR SMART; SM00248; ANK; 6.
 DR PROSITE; PS0088; ANK_REPEAT; 4.
 DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
 KW ANK repeat; Hypothetical protein.
 SQ SEQUENCE 823 AA; 93968 MW; 9D2DD5F52C379B2 CRC64;

Query Match 7.7%; Score 396; DB 2; Length 823;
 Best Local Similarity 17.6%; Pred. No. 7.2e-08;
 Matches 178; Conservative 100; Mismatches 155; Indels 580; Gaps 18;

Db 1330 KKAEEKKADELKKSEKKKKADELKKSEKKKADE-----LKKAAEEKKKADELKKAAEE 1384
Qy 597 QKEPDTLSGKLESPDKGLLKPTCGMKISLPNKALELKDOR--ETFFKAEDVSSVESTFSL 654
Db 1385 KKADELKKAAEEKKKADELKK-----KAEKKKADELKKAAEEKKKADELKKAAEE 1436
Qy 655 FGKPTTENSQSTKVBEEDFNLTATKEGATKTVTGQQRDGIITIERAPQDOTNKMPTSELGRK 714
Db 1437 --KKAENLK--KAE-----KKKADELKKAAEEKKKADELKKAAEE-----KKKADELKK 1484
Qy 715 EDTKSTSDSEIISVSDTQNYECLPRTATYQKEIKTTNGKIEBSPKPSHFHPATQMNSVP 774
Db 1485 AEEKKKAD-ELKAAEEKKKADELKKAAEEKK--KADELKKAAEEKKKADELKKAAEEKKAAEE 1541
Qy 775 NKGLWKNQTLRADSTTLKILDALPSCERGR-----ELKKNDCQEI 817
Db 1542 KKKVEQKKREBERNNAL--RAEILKQIEKKRIEVMKLYEEKKKMAEQKKKEEBEKI 1599
Qy 818 TA-----KMEQMKNFVQLQKELSEAKETIKSQLENQKAKWEQELCSVRLPLNQE 866
Db 1600 KAEQLKKEEBEKKKVEQLKKK--EEBKKKAEQLKKEEBEENKIKAEQ-----LKKKEE 1650
Qy 867 EKKRNVDILK--EKIRPEEQRLKLEVKHQLEQTLRIQDIELKSVTSNLNQVSHTHES 923
Db 1651 BEKKAAELKKEEBEKKKAEQLKKEEBEKKKVEQLKKEEBEKKKAAE-----1698
Qy 924 ENDLPHENCMKKETAMLKLEVATLKHQHQVKENKYFEDIKLOEKNAELQMTLKLKQKT 983
Db 1699 -----QKKKEEBEENKIKVEQLKKEEB--EKKKAAELKKEEBEKKKVEQLKKEEB-- 1746
Qy 984 VTKRASQYREQLKVLTAE 1001
Db 1747 --KKAEEIRKEKEAVIEE 1762

RESULT 12
Q7K5Q5 PRELIMINARY; PRT; 2019 AA.
AC Q7K5Q5;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Erythrocyte binding protein 2.
GN Name=maebi;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_Taxid=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22072559; PubMed=1206768; DOI=10.1016/S0166-6851(02)00067-1;
RA Blair P.L., Kappe S.H., Maciel J.E., Balu D.J., Adams J.H.;
RT "Plasmodium falciparum MAEBL is a unique member of the ebl family.";
RL Mol. Biochem. Parasitol. 122:35-44(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22077637; PubMed=12082132;
RA Michon P., Stevens J.R., Kaneko O., Adams J.H.;
RT "Evolutionary relationships of conserved cysteine-rich motifs in
adhesive molecules of malaria parasites.";
RL Mol. Biol. Evol. 19:1128-1142(2002).
RN [3]
RP SEQUENCE FROM N.A.
RA Blair P.L., Kappe S.H.I., Maciel J.E., Balu D.J., Adams B.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY042084; AAQ73469.1; -
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
SQ SEQUENCE 2019 AA; 239410 MW; A626F2C684C08785 CRC64;

Query Match 6.9%; Score 359; DB 2; Length 2019;
Best Local Similarity 22.4%; Pred. No. 5.6e-06;
Matches 233; Conservative 159; Mismatches 398; Indels 248; Gaps 42;

Qy 109 RKLPNQNTNPECTGTCTPDEAPLAERTPDTAESLLEKTPDEARLVEGTSAKIOCLG 168
Db 828 RNLPAVYHLTNECVILGTHEQ-----ERTNSCRRTKEE-----KKPNCQILR 871
Qy 169 KATSGK---EQSTTEETPRKILRPTKETSEKFSWPAKERSRKI---TWEEK- 213
Db 872 KTTDSKDMTWYSS-----FIRPDYETKCPPRYPLKS---KVFGTFDQKTGCKSLMDK 921
Qy 214 -----ETSVKTCVAGVTPNKTEVLEKGTSMIACTKETSTKASTNVDDVS----- 260
Db 922 AYEVGINKFSVCLYLEFLVSPK--DLYNSGRNNTWGIWAADHSVN--ENNIETANGKCVHL 978
Qy 261 -VEPIFLSGFTRITIENSQCTKVBEEDFNLTATKIISKSAQNYTCLPDATYQKDIKT----- 314
Db 979 VVKPTCVIDKENHFSFTALTANTVDFNQSVNIRKIEELTEYGNNDVVLKEKEINNEPIDN 1038
Qy 315 -----INH-----IEDQMPFSESKREDEEYSDWSGSLFESSAKTOV--CIP 355
Db 1039 VNETHKINRSHVNSMERKNKFSYKENEDYDQEKVDETYSEFG-LFEEARKTGTGRIEE 1097
Qy 356 ESMYQKVM-----EINREVE-----E 371
Db 1098 ESKKEAMKRAEDARRITEARRAEDARRIEARRAEDARRVEIARRVEDARRIEISRAE 1157
Qy 372 LPEKPSAFKPAVENQKTPVKNKAFELKNEQTLRAAQMPFSESKQKDDSENSWDSSEPCETV 431
Db 1158 DAKRIEARRAIEVRRAEALRKA---EDARRIEAARRY--ENERRIEARRVEDEKRIEAV 1212
Qy 432 S-----QKDVYLPKATQKQEPDTLSGKLESPV-----KQGLLKPTGCRKVSLPNKALEL 481
Db 1213 KRAEEVRKDBEEAKRAEKERNNEIRKFEEMAHFARRQAAIKAEKRADELLKKAEBK 1272
Qy 482 KDRETFKAESPDK--DGLLKPTCGRKVS--LPNKALELKORETLKAESPNDGGLKPTCG 537
Db 1273 KKADELKKSEKKKADELKKAAEEKKKADELKKAAEEKKKADELKKAAEEKK--KADEV 1329
Qy 538 KVSPLNKALEL--KDRETFKAAQMPFSESKQKDDSENSWDFSEFLETLQNDVCLPKATH 596
Db 1330 KKAEEKKKADELKKSEKKKADELKKSEKKKADE-----LKKAAEEKKKADELKKAAEE 1384
Qy 597 QKEPDTLSGKLESPDKGLLKPTCGMKISLPNKALELKDOR--ETFFKAEDVSSVESTFSL 654
Db 1385 KKADELKKAAEEKKKADELKK-----KAEKKKADELKKAAEEKKKADELKKAAEE--- 1436
Qy 655 FGKPTTENSQSTKVBEEDFNLTATKEGATKTVTGQQRDGIITIERAPQDOTNKMPTSELGRK 714
Db 1437 --KKAENLK--KAE-----KKKADELKKAAEEKKKADELKKAAEE-----KKKADELKK 1484
Qy 715 EDTKSTSDSEIISVSDTQNYECLPRTATYQKEIKTTNGKIEBSPKPSHFHPATQMNSVP 774
Db 1485 AEEKKKAD-ELKAAEEKKKADELKKAAEEKK--KADELKKAAEEKKKADELKKAAEEKKAAEE 1541
Qy 775 NKGLWKNQTLRADSTTLKILDALPSCERGR-----ELKKNDCQEI 817
Db 1542 KKKVEQKKREBERNNAL--RAEILKQIEKKRIEVMKLYEEKKKMAEQKKKEEBEKI 1599
Qy 818 TA-----KMEQMKNFVQLQKELSEAKETIKSQLENQKAKWEQELCSVRLPLNQE 866
Db 1600 KAEQLKKEEBEKKKVEQLKKK--EEBKKKAEQLKKEEBEENKIKAEQ-----LKKKEE 1650
Qy 867 EKKRNVDILK--EKIRPEEQRLKLEVKHQLEQTLRIQDIELKSVTSNLNQVSHTHES 923
Db 1651 EKKKAAELKKEEBEKKKAEQLKKEEBEKKKVEQLKKEEBEKKKAAE-----1698
Qy 924 ENDLPHENCMKKETAMLKLEVATLKHQHQVKENKYFEDIKLOEKNAELQMTLKLKQKT 983
Db 1699 -----QKKKEEBEENKIKVEQLKKEEB--EKKKAAELKKEEBEKKKVEQLKKEEB-- 1746
Qy 984 VTKRASQYREQLKVLTAE 1001
Db 1747 --KKAEEIRKEKEAVIEE 1762

RA Ghai M., Dutta S., Hall T., Freilich D., Ockenhouse C.;
RT "Identification, expression, and functional characterization of MAEBL,
RT a sporozoite and asexual blood stage chimeric erythrocyte-binding
RL protein of Plasmodium falciparum";
DR Mol. Biochem. Parasitol. 123:35-45 (2002).
DR EMBL; AF042084; AAL10509.1; -
DR EMBL; AF400002; AAM90625.1; -
DR HSP; P04368; IIC2.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
SQ SEQUENCE 2055 AA; 243211 MW; 783F48BB7286B1F3 CRC64;

Query Match 6.9%; Score 359; DB 2; Length 2055;
Best Local Similarity 22.4%; Pred. No. 5.7e-06;
Matches 233; Conservative 159; Mismatches 398; Indels 248; Gaps 42;

QY 109 RKLKPNPONTNPEGSTGTGTPDEAALPABRTPTDAESLLEKTPDEAARLVEGTSAKIQCIG 168
DB 828 RNLPAYNHLNNECVILTHQ-----ERTNSCRTKEE-----KKPNCQILR 871
QY 169 KATSGK---FQSTETPRKILRPTKETSEKFSWPAKERSRKI--TWEEK----- 213
DB 872 KTTDSKOWTVSS-----FIRPDYETKPPRYPLKS--KVFGTDPQTKGCKSLMDK 921
QY 214 -----ETSVKTECVAGVTNKTETVLEKTSNMIACPTKETSTYKASTNVDS-- 260
DB 922 AYEVGINKFSFLEYLFLVSPK--DLYNSGRNYYWGIWAADHSVN--ENNIEANGKCYHL 978
QY 261 -VEPIFSFGRTIENSQCTVEEDFNATKIISKSAQNTCLPDATYQDKT----- 314
DB 979 VVKTCVIDKENHFSFALTANTVFNOSVIRKIEELTEGNDDVLKEIINNEPIDN 1038
QY 315 -----INHK-----IEDQMFPESEKREDEEYSDSGSLFESSAKTV--CIP 355
DB 1039 VNETKINKSHVSMERNKPSYKENEYDQMEKNVEDYSEFG-LFEEARKTETGRIEE 1097
QY 356 ESMYQKVM-----EINREVE-----E 371
DB 1098 ESKKAEAKRAEDARRIEBEARRAEDARRIEARRVEDARRIEISRAE 1157
QY 372 LPEKPSAFKPAVEMOKTPNKAFLKNEQTIRAAQMPFSESKOKDEENSDSPCSTV 431
DB 1158 DAKRIEARRAIEVRAELRKA--EDARRIEARRY--ENERRIEARRYEDEKRIEAV 1212
QY 432 S-----QKDVLPKATHQKEFDTLGKLEESPV-----KGLLKPTCGRKVSLPNKALEL 481
DB 1213 KRAEVRKDEEBAKAEKERNNEEIRKFEARMAMHARRQAAIKAEKKKADLKKAEK 1272
QY 482 KORETFKAESPK--DGLLKPTCGKRV--LPNKALELKDRETLKAESPDNDGLLKPTCG 537
DB 1273 KKADELKSEKKKADLKKAEKKKADLKKAEKKKADLKKAEKK--KADEV 1329
QY 538 RKVSLPNKALEL-KDRETFKAQMPFSESKOKDEENSDPFLETLLQNDVCLPKATH 596
DB 1330 KKAEEKKKADLKKSEKKKADLKKSEKKKADLKKSEKKKADLKKAEK-----LKKAEKKKADLKKAE 1384
QY 597 QKEFTTLGKLEESPDGLLKPTCGMKISLPNKALELKD--ETTFKAEDYSSVESTFSL 654
DB 1385 KKADELKKAEEKKKADLKK-----KAEKKKADLKKAEKKKADLKKAEK--- 1436
QY 655 FGKPTTENSQSTKVEDENLITTKGATVTVGQQRDGIIERAPQDQTNKMPSELORK 714
DB 1437 --KKAENLK--KAE-----KKKADLKKAEKKKADLKKAEK-----KKKADLKKK 1484
QY 715 EDTKSTSEIISVSDTQNYECLPATYQKEIKTTNGKIESPEKSPHFEPATQMNSVP 774
DB 1485 ABEKKKAD-ELKAEKKKADLKKAEKK--KADLKKAEKKKADLKKAEKKKAE 1541
QY 775 NKGLEWKNQTLRADSTTLKSLDALPSCGR-----ELKKDNCRQI 817

DB 1542 KKKVEQKKREBERNNMAL--RRAILKQIEKKRIEYVMKLYEEKKMAEQLKKEEKEKI 1599
QY 818 TA-----KWEQMKKFCVLQKELSEAKEIKSOLENQKAKWQELCSVRLPLNOE 866
DB 1600 KAEQKKKEEKKVQMLKKK--ESEEKKAEQKKKEEENKIKASQ-----LKKKEE 1650
QY 867 EEKERNVDILK--EKIRPEEQKLEKLVKHQLEQTILRIQDIELKSVTSLNQVSHTHES 923
DB 1651 EEEKKAEELKKEEKEEKKAEQKKKEEKKVQELKKKEEKKKAE----- 1698
QY 924 ENDLFHENCMLKKEIAMLKLEVATLKHQVKENKYFEDIKLOEKNAELOMTLKQKT 983
DB 1699 -----QLKKEEENKIKVEQKKKEE--EEKKAEELKKEEKKVQVQKKKEE-- 1746
QY 984 VTKEASQYREQLKVLTAE 1001
DB 1747 --KKAEEIRKEKEAVIEE 1762

RESULT 15
Q8ISFS PRELIMINARY; PRT: 10578 AA.
ID Q8ISFS
AC Q8ISFS;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE lmda_1 protein.
GN Name=iso1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22269627; PubMed=12381307; DOI=10.1016/S0022-2836(02)00970-1;
RA Flaherty D., Gernert K., Shmeleva N., Tang X., Mercer K.,
RA Borodovsky M., Benlan G.;
RT "Titins in Caenorhabditis elegans with Unusual Features: Coiled-coil
RT Domains, Novel Regulation of Kinase Activity and Two New Possible
RT Elastic Regions";
RL J. Mol. Biol. 323:533-549 (2002).
DR EMBL; AY130758; AAM61519.1; -
DR PIR; E89066; E89066.
DR PIR; T33247; T33247.
DR HSSP; Q10466; lBPV.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR008941; TPR-like.
DR InterPro; IPR011046; WD40-like.
DR Pfam; PF00041; fn3; 2.
DR Pfam; PF00041; fn3; 3.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00408; IGC2; 3.
DR PROSITE; PS0853; FN3; 2.
DR PROSITE; PS0835; IG_LIKE; 11.
SQ SEQUENCE 10578 AA; 1174295 MW; 71BF50C502FB4F16 CRC64;

Query Match 6.9%; Score 357.5; DB 2; Length 10578;
Best Local Similarity 21.7%; Pred. No. 4.2e-05;
Matches 269; Conservative 157; Mismatches 454; Indels 357; Gaps 47;

QY 36 EFLTKNANAFNESKCTALMLAIC-----EGSSEIVG--MLQQNVDPVPAEDIHGITA 88
DB 8993 EFYAFRIVAVNGFEGAPSEIIIEVNTLDYDQESFDFAGEBELKLDVQVNVNEVTEIT 9052
QY 89 BRYAARGVNYHQLEHRLKLPKNQNTNPEGSTGTGTPDEAALPABRTPTDAESLLEK 148
DB 9053 EE-----SEVTEEHRLKKKKSKK--KTTDEPDLDEIALEVSSDITSSLEI 9099
QY 149 T-----PDEARLVEGTSAKTICLGKATSGKFEQSTETPRKILRPTKETSEKFSWPAKE 203

1115 TUGT BLINK (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:56:15 ; Search time 28.5932 Seconds
(without alignments)
3402.035 Million cell updates/sec

Title: US-09-602-362E-27
Perfect score: 5173
Sequence: 1 MVATLLSYGAVIEVQNKASL.....REQLKVLTAENTMLTSLKLE 1011

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	341.5	6.6	1427	2 S22695	restin - human
2	340.5	6.6	1790	2 S67593	transport protein
3	337	6.5	1392	2 A4336	microtubule-vesicl
4	321.5	6.2	1938	2 JC5421	smooth muscle myos
5	321.5	6.2	1972	2 JC5420	smooth muscle myos
6	321	6.2	2663	1 S28261	centromere protein
7	318.5	6.2	2116	2 A26655	myosin heavy chain
8	315	6.1	1972	1 A41604	myosin heavy chain
9	315	6.1	1488	2 T34418	hypothetical prote
10	312.5	6.0	5327	2 T13564	microtubule-associ
11	311	6.0	3259	1 A56539	giantin - human
12	310	6.0	1961	1 A61231	myosin heavy chain
13	310	6.0	3225	2 I52300	giantin - human
14	307.5	5.9	2954	2 T14156	kinesin-related pr
15	302.5	5.8	3187	2 JC5837	364K Golgi complex
16	300.5	5.8	1313	2 F96673	hypothetical prote
17	300.5	5.8	2253	2 T30336	nuclear/mitotic ap
18	300	5.8	1979	1 S03166	myosin heavy chain
19	295.5	5.7	2057	2 S61477	myosin II heavy ch
20	294.5	5.7	2017	1 A36014	myosin heavy chain
21	293.5	5.7	1957	2 T38077	hypothetical colle
22	293	5.7	1939	2 T18372	repeat organellar
23	292.5	5.7	1410	1 A57013	early endosome ant
24	291.5	5.6	2677	2 A38194	desmoplakin I - hu
25	290	5.6	2139	2 T18296	myosin heavy chain
26	287.5	5.6	2331	2 T25410	hypothetical prote
27	286	5.5	993	2 S49461	synaptonemal compl
28	285	5.5	1999	1 S21801	myosin heavy chain
29	284.5	5.5	1538	2 T29095	cardiac muscle fac

RESULT 1
S22695
restin - human
C:Species: Homo sapiens (man)
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
C:Accession: S22695; S19853
R:Billbe, G.; Delabie, J.; Brueggen, J.; Richener, H.; Asselbergs, F.A.M.; Cerletti, N.;
EMBO J. 11, 2103-2113, 1992
A:Title: Restin: a novel intermediate filament-associated protein highly expressed in th
A:Reference number: S22695; MUID:92289675; PMID:1600942
A:Accession: S22695
A:Molecule type: mRNA
A:Residues: 1-1427 <BIL>
A:Cross-references: UNIPROT:P30622; EMBL:X64038; NID:g35998; PIDN:CAA46050.1; PID:g35999
C:Keywords: cytoskeleton

Query Match 6.6%; Score 341.5; DB 2; Length 1427;
Best Local Similarity 22.4%; Pred. No. 2.5e-07;
Matches 251; Conservative 185; Mismatches 414; Indels 271; Gaps 54;

Qy 1 MVATLLSYGAVIEVQNKASLTPLLAIQKRSKQVTEFLLTKNANANAFNSKCTALMLAI 60
Db 300 MATTSASLKRSPSASSLSMSSVASSVSSRPSRT--GLLTSTSRYARKISGTTALQAL 357
Qy 61 CEGSSEIVGMLQONVDVFAEDIHGITAERYAARGVNYIHQOLL-----EHRKLPK 113
Db 358 KEKQHIETQLLAERDLE-----RAEVAKATSHVGEIEQELALARDGHDQHVLEL-- 406
Qy 114 NPQNTNPEGTSTGTPDEAAPLAERTPDATBSLLEKTPDEAARLVEGTSAKIQCLGKATSG 173
Db 407 -----EAGWDQLRTVVEAADREKVELLNQL--EEKKRVEDLQFRVE--ESITKG 453
Qy 174 KFEOSTETPRKILRPTKSEKFSWPAKERSKRIWTBEEKTSVKTECVAGVTNKTVL 233
Db 454 DLETQTKLEHARI----KELEQSLLF--EKTAKDLQRELEDTRVAT-----VSEKSRIM 502
Qy 234 EKGTSNMIACTPKET-----STKASTNVD--VSSVEPIFSLFG----TRTIENSQCT 279
Db 503 E--LEKDALRVQVAEIRRRLESNKPGVDMSLSLLQELISSLOEKLEVTRTDHQREIT 560
Qy 280 KVEEDFNLATKIISKAAQNYTCLPDATYQDKI---TINHKIEDQMPFSPSKREDEBY 336
Db 561 SLKEHFGAR-----EETHQKEIKALYTATEKLSKENESLSKLEHANKE 604
Qy 337 SWDGSGLFESSAKTQVCIPESMYQKME-----INREVEIPE-KPAFAFPAVEMQ 386
Db 605 NSDVIALWKSLETAI----ASHQQAAMEELKVSFSGKLGTTABFAELKTQIKMRLDYQ 660
Qy 387 KTVPN-----KAFELKNEQTLRAQMFPSESKQDKDEENSWDESPCETVSKQVYL 438
Db 661 HETENLQNDQDSRAAHAKMEALRAKLM-----KVIKEKENS-----LEAIRSK---L 706

ALIGNMENTS

Qy 439 PKA--THQKEDFTLGGKLESPVKD---GLLKPTCGRKVS-LPNKALELKDRETFKAESP 492
 Db 707 DKAEDQHLVEMEDTLNKLEAEIKVKELEVLQAKCNEQTKVIDNFTSQLKATE-----759
 Qy 493 KDKGLLKPTCGRKVS-LPNKALELKDRETFK-----AESPDNDGLLKPTCGR 538
 Db 760 --EXLLDLDALRKASSSEKMKLRQOQEAABEVKQIKHLEIEKNAESKASSITRELOGR 817
 Qy 539 KVS-LPNKALELKD---RETFKAAQMPFSESKQKDEENSWDFSEFLETLNDVCLPKA 594
 Db 818 ELKLTNLQENLSEVQKTELEKQLKLEKFAEABEAVSVQSMQETV-----NK 869
 Qy 595 THQKE--PDTLSGKLESPDKGLLKPTCGMKISLPNKALELKDRETFKAEDVSSVSTF 652
 Db 870 LHQKEQFNMLSSDLK-----LRENADMEAKFREK--DERBEQLIKAKE--KLNDI 919
 Qy 653 SLFGKPTTEN--SOSTKVEEDFNLTTKGATKTVTGQOERDI-----GIER 697
 Db 920 AEIMKMGDNGSOLTKMDELRL-----KERDVEELQLKLTKANENASFLQK 966
 Qy 698 APQDQTNKMTS-----ELGRKEDTKSTDSSEIISVSOTNYVECLP-BATYOK---E 745
 Db 967 SIEDMTVKAESQOQEAABKHEEKELERKLSLE--KKMETSHNQCOELKARYERATSE 1024
 Qy 746 IKTTNGKLEESPEK-----PSHFEPATEMOSVNPKNKLEWKNKOTLRADSTTLKILDAL 800
 Db 1025 TKTRHEEILQNLQKTLTDTEDKLGARE--ENSGLLQLELEELRKQADKAAQTAIE--DAM 1081
 Qy 801 PSCERGERLKKDNCBQITAKMBQMKNFQVLOKELSEAKE--IKSQLENQAKWEQELCS 858
 Db 1082 QIME---QWTKETETL-ASLEDTQTNAKLQNELDTLKENNLKNVEELNKS--ELLT 1134
 Qy 859 VRLPLNQE--EKRNRVNDILK-----EKRPEBOLARK--889
 Db 1135 VE---NQKMEFRKEIETLLQAAQKSOQSLAQLQENVKLABELGRSDEVTSHQKLEEE 1191
 Qy 890 -----LEVGHOLEOTLRIODIELKSVTSNLNVSHTHESENDFLPHENCMKKKEIAML 941
 Db 1192 RVLNQLLEMKRKSFKIDADEKASLOKIS-ITSALLTEKDAELEK--LRNEVTVL 1248
 Qy 942 KLEVALTKHQHOKENKYPEDIKI-IQENKAEQLMTLKLQ 981
 Db 1249 RGENASAKLSHVSQVLTESDKVKLELKVQNLQLEKQENKQ 1289
 RESULT 2
 S67593
 transport protein USO1 - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein D2552; protein YDL058w
 C:Species: Saccharomyces cerevisiae
 C>Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
 C:Accession: S67593; A38455; S30782
 R:Blocher, H.; Brandt, P.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S67587
 A:Accession: S67593
 A:Molecule type: DNA
 A:Residues: 1-1790 <BL0>
 A:Cross-references: UNIPROT:Q07380; EMBL:Z74106; NID:G1431058; PID:e253003; PID:G1431059
 A:Experimental source: strain S288C
 R:Nakajima, H.; Hirata, A.; Ogawa, Y.; Yonehara, T.; Yoda, K.; Yamasaki, M.
 J. Cell Biol. 113, 245-260, 1991
 A:Title: A cytoskeleton-related gene, USO1, is required for intracellular protein transp
 A:Reference number: A38455; MUID:91185402; PMID:2010462
 A:Accession: A38455
 A:Molecule type: DNA
 A:Residues: 1-389, 'TA', 392-724, 'S', 726-1790 <NAK>
 A:Cross-references: GB:X54378; NID:G4777; PIDN:CAA38253.1; PID:G4778
 A:Note: the authors translated the codon ACT for residue 768 as Ile
 R:Hostetter, M.K.; Herman, D.J.; Bendel, C.M.; McClellan, M.; Tao, N.; Kendrick, K.E.
 submitted to the EMBL Data Library, February 1993
 A:Description: An integrin analogue in Saccharomyces cerevisiae.
 A:Reference number: S30782

A:Accession: S30782
 A:Molecule type: DNA
 A:Residues: 71-846, 'E', 848-923, 'K', 925-1252, 'I', 1254-1318, 'V', 1320-1460, 'S', 1462-1580, 'S'
 A:Cross-references: EMBL:L03188
 C:Genetics:
 A:Gene: SGD:USO1; INT1
 A:Cross-references: SGD:S0002216; MIPS:YDL058w
 A:Map position: 4L
 C:Keywords: Coiled coil; transmembrane protein
 F:326-342/Domain: transmembrane #status predicted <TM1>
 F:394-410/Domain: transmembrane #status predicted <TM2>
 F:617-633/Domain: transmembrane #status predicted <TM3>
 Query Match 6.6%; Score 340.5; DB 2; Length 1790;
 Best Local Similarity 20.7%; Pred. No. 3.6e-07;
 Matches 250; Conservative 185; Mismatches 412; Indels 359; Gaps 53;
 Qy 1 MVATLSYGVAVI--EVQNKASITPLL-LAIQKRSQT-----VEFL---LTKNANAN 46
 Db 601 VIKSLISFSYQIQEDVTKLVTLMLGVAYEFSSESPPRKEYFEFTIKTLGKONYAS 660
 Qy 47 AFNRSCTALMLAICESSSEIVGMLQQN-----VDVFAEDIHGITAERVAAR 95
 Db 661 RIKQFKKDSYFSKVDNEDSILPELDETGLPKVYFTFIQLPNENIYRI---RTALSH 717
 Qy 96 GVNYY--IHQQLLEHRLKLPKNQNTPEGTSTGTPDEAALERTPDPAESLLEK---TP 150
 Db 718 DPDEEPINKISPEVEKLQR--QCTKLKG-----EITSLOTETESTHENLTKLIALT 768
 Qy 151 DEARLVGTSKAIQICLGKATSKKFEQ--STEETPRK-----ILRPTKSTSEKFSW 199
 Db 769 NEHKELDE---KQIILNSSHSLKENFSILELKNVRLSDLEMTQRLDVLTKDKENQ 824
 Qy 200 PAKERSKITWEKETSVKTECVAGVTPNKTVELEKGTSNMIACPTKETSTKASTNVDS 259
 Db 825 TAL-LEKSTIHKQEDSIKT-----LEKLETLIS-----QKKKAEDGINKM 865
 Qy 260 SVEPIFLSGFTRTIENSQCTKVEEDFNLA-----TKIISKSAQAQNYTCLPDATYQKDI 312
 Db 866 G-KDLPALSRQMAVEENCKNLQKEDKSNVNHQKETSLSKLEADIAAKIT-----EI 915
 Qy 313 KTIHKLQDMQFSPSEKREDEEYSDSGSILFESSAKTQVCIPESMYQKWEINREVEEL 372
 Db 916 KAINENLEEMKIQCNNLSKKEHIS-----KEL 943
 Qy 373 PEKPSAFK-----PAVEMQKTVPN--KAFELKNEQTLRAAQMPFSEK--QKDDENSW 422
 Db 944 VEYKSRFQSHDNLVAKLTKLSLANNYKDMQAEENSLIKAVESKNESLQSLNQLNKI 1003
 Qy 423 DSESFCETVSKQVYLPKATHQKEFDLTSGLKLESPVKDGLLKPTCGRKVS-LPNKALEL 482
 Db 1004 DS-----MSQE-----KENFQIERGSIENIEQ-----LKKT-----IS 1032
 Qy 483 DRETFKAESPDGGLLKPTCGRKVS-LPNKALE-----LKDRETLKAESP 526
 Db 1033 DLEQTKEEIISKSDSSKDEYESQISLKEKLETATTANDENVNKISELTKEEELAEALA 1092
 Qy 527 DNDGLLKPTCGRKVS-LPNKAL-ELKDRETFKAAQMPFSESKQKDDENSWDFSEFLETL 585
 Db 1093 AYKN-LKNELETLETSEKLEKEVKNBEHLKBEKIQLKEKATETKQOLNLSRANLESL- 1150
 Qy 586 QNDVCLPKATHQKEFDLTSGLKLESPDKDGLLKPTCGMKISLPNKALELKDRETFKAEDV 645
 Db 1151 -----EKHEHDAAQQLKYEE-----QIANKEQYNEEI 1179
 Qy 646 SSVSTFSLFGKPTTENSQTKVEEDFNLTTKGATKTVTGQOERDIGIIERAPQDQPNK 705
 Db 1180 SOLNDEIT-----STQOENESIKKND-----ELEGEVKAAMKSTSEQSN-LKKSEIDALN- 1229
 Qy 706 MPTSELGRKEDTKSTG-----DSEIISVSDTQNYECLPEATYQKEIKTINGKIESP 757
 Db 1230 LQIKELKKKNETNEASLLESIKSVESVTKIKELQD-EC---NFKKEVSELEDKKASE 1285

```
Qy 758 EKSPHFPATPMQNSVFNKGLWKNKOTLRADSTTLTKILDALPSCRGRE-----L 809
Db 1286 DKNSKY---LELOKE-----SEKIBELDAKTTELKIQELITNLKAKESSELSRL 1336
Qy 810 KKNCEQ---ITAKMEQMKKFCV-----LOKELSE----- 837
Db 1337 KTSSEERKNAEOLKELKNEIQKQAFKPKLLNKGSGSTITTOEYSEKINTLEDELIR 1396
Qy 838 -----AKEI---KSQLE-----NOKAKWEQELCSVRLPLNQBEERK--- 869
Db 1397 LQENELKAKEIDNTRSELEKVSLSNDELLEEKQNTIKSLQDEILSYKDKITRNDEKLLS 1456
Qy 870 -----RRNVILKPKIRPE-----QURKKLVKHQLEQTLRI 902
Db 1457 IERNKRDLSLKBQLRAAQESKAKVBEGLKLBEESSKKAELKSKEMMKKLESTIES 1516
Qy 903 QDELKSVTSLNQVSH-----THESENDLPHEMCKKETAAMLKLEVATLKHQHV 954
Db 1517 NETELKSMETIRKSDKLFOSKKSABEEDIKNLOHESDLSIRNESEKIDIELKSLRI 1576
Qy 955 --KENKYFEDIK---ILQEK---NAELQMTLKLKQTVTKRASQYREQKVLTAENTML 1005
Db 1577 EAKSGSELETVKQELNNAQEKIRINAEENTVLSKLEDIERELKDKQAEIKSNQBEKELL 1636
Qy 1006 TSKLKE 1011
Db 1637 TSLRKE 1642

RESULT 3
microtubule-vesicle linker CLIP-170 - human
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Nov-1999
C;Accession: A43336
R;Pierre, P.; Scheel, J.; Rickard, J.E.; Kreis, T.E.
Cell 70, 887-900, 1992
A;Title: CLIP-170 links endocytic vesicles to microtubules.
A;Reference number: A43336; MUID:92405160; PMID:1356075
A;Accession: A43336
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1392 <PIE>
A;Cross-references: GB:M97501; NID:g180621; PID:AAA35693.1; PID:g180622

Query Match 6.58; Score 337; DB 2; Length 1392;
Best Local Similarity 22.28; Pred. No. 3.8e-07;
Matches 246; Conservative 181; Mismatches 401; Indels 280; Gaps 52;

Qy 1 MVAITLLSYGAVIEQNKASLTPLLLAIQKRSKOTVEFLITKNANANAFNESCCTALMLAI 60
Db 300 MATTSASLKRSPSASSLSMSSVASSVSSRPSRT--GLLTETSSRYARKISGTTALQAL 357
Qy 61 CEGSSEIVGMLLQNVDFVFAEDIHGITAERYAAAGVNYHQHLLHIRKLPRNPQNTNP 120
Db 358 KEKQOHTIEQLAERDL-----ERAEVAKATSHV----- 385
Qy 121 EGTSTGTGPDENAPLAERTPTDAESLLEKTPDEARLVETGS-AKIQCLGKATSGKFEQST 179
Db 386 -----GEIQLALARDGHQHVLELAKMDQLFTWVEAADREKVELLNQ-----L 431
Qy 180 EETPRKI---LRPTKETSEKFSW---PAKERSRKITWEEKTSVKTECVAGVTPNKTEV 232
Db 432 EEEKRKVEDLQFVEESITKGDLEVATVSEKSR-IMELEKDALARVQEVA-----EL 483
Qy 233 LEKGTSMIACTKETSTKASTNVVDSSVEPIFSLFGTRTIENSQCCTKVEEDFNATKII 292
Db 484 RRRLESNK---PAGDVMSLSLLQEISSLOE--KLEVTRTDHOREITSKEHFCAK---- 534
Qy 293 SKSAAQNYTCLPDATYOKDIK---TINHKTEDQMPFSPKKEEEDSVSWDSGSFPSSAK 349
Db 535 -----EETHQKEIKALYATATEKLSKENESLKSLEHANKENSVDIALWKSLE 582
```

```
Qy 350 TOVCIPESMYQKQVME-----INREVELPE-KPSAFKPAVEMQKTVPN----- 391
Db 583 TAI-----ASHQQAAMEELKVSFKGLGTETAFABELKTIQIEKWLDTQHEIENLQOQDSE 638
Qy 392 KAPELKNEQTLRAAQMPFSESOKKDDSENSWDSSEPCETVSQKDVYLPKA--THQKEPDT 449
Db 639 RAAHAKEMEARAKLM-----KVTKERENS-----LEAIRSK---LDKABDQHLVEMED 684
Qy 450 LSGKLBSPPVKD---GLLKPTCGRKVS-LPNKALELKDRETFKAESPDKGLLKPTCGRK 505
Db 685 TLNKLQAEAIKVKLEVLQAKCNEQTKVIDNFTSOLKATE-----EKLDDLALRK 735
Qy 506 VSLPNKALELKDRETLK-----AESPDNDGLLKPTCGRKVSIPNKALEKLD 551
Db 736 ASSEKSEMKKLQOLQAABKQIKHLEIEKNAESSKASSITRELQGRKELKLTNLQENLSE 795
Qy 552 ----RETFKAAQMPFSESOKKDDSENSWDFESFLETLILLONDVCLPKATHOKE--PDTLSG 605
Db 796 VSQVKETLEKLOLILKEKFAEASEEAVSVQKSMQETV-----NKLHQKEEQFNMLSS 847
Qy 606 KLBSPPDKDGLLKPTCGMKISLPNKALELKDRETFKAEDVSSVSESTFSLFGKPTTEN-SQ 664
Db 848 DLEK-----LRENLMADMEAKFREK--DEREEQLIKAKE--KLENDIAEIMKMSGDSNSQ 897
Qy 665 STKVEEDFNLTKEGATKTVTGOQERDI-----GIIRAPQDQTNMPTS- 709
Db 898 LTKMNDLRL-----KERDVEELQLKLTKANENASFLQKSIEDMTVKAEOQS 944
Qy 710 -----ELGRKEDTKSTDSEIISVSDTQNYECLP-EATYQK---ELKTTNGKLEESPE 758
Db 945 QEAAKGEHEEKKELEKRLSDLE--KQMETSHNOCOELKARYATSETTKTKEHSEILQNLQ 1002
Qy 759 K-----PSHPEPATPMQNSVFNKGLWKNKOTLRADSTTLTKILDALPSCRGRELKKN 813
Db 1003 KTLTLDTEKLGARE-ENSGLLQLEBLRQAEAKAAQTAE--DAMQIME---QMTKEK 1056
Qy 814 CEQITAKMEQMKKFCVLQKELSEAKE--IKSLENOKAKWEQELCSVRLPLNQBEERK 870
Db 1057 TETL-ASLEDTKQTNAKLQNELDTLKNNLKNVVELNKS--ELLTVE---NQMSEBFR 1109
Qy 871 RNVLDILK-----EKIRPEEQLRKK-----LEVKH 894
Db 1110 KEIETLKQAAQKSSQSLALQEBENVKLAELGSRDEVTSHQKLEEBERSVLNNQLEMKK 1169
Qy 895 QLEQTLRIQIELKSVTSLNNOVSHSTHESENDLPHEMCKKETAAMLKLEVATLKHQHV 954
Db 1170 RESKFIKDADEEKASLOKSIS-ITSALLTEKDALEK--LRNEVTVLGENASAKSLHSV 1226
Qy 955 KENKYFEDIKI-LOEKNAELQMTLKLKQ 981
Db 1227 VOTLESQVKLELKVKNLELQLENKQ 1254

RESULT 4
JC5421
smooth muscle myosin heavy chain 2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 09-Jul-2004
C;Accession: JC5421
R;Hasegawa, K.; Arakawa, E.; Oda, S.; Matsuda, Y.
Biochem. Biophys. Res. Commun. 232, 313-316, 1997
A;Title: Molecular cloning and expression of murine smooth muscle myosin heavy chains.
A;Reference number: JC5420; MUID:97242182; PMID:9125171
A;Accession: JC5421
A;Molecule type: mRNA
A;Residues: 1-1938 <HAS>
A;Cross-references: UNIPROT:O08638; DDBJ:D85924; NID:g1945079; PID:BAAL1691.1; PID:g194
A;Experimental source: smooth muscle
C;Comment: This protein plays a role in smooth muscle cell contraction.
C;Superfamily: myosin heavy chain; myosin motor domain homology
C;Keywords: nucleotide binding; P-loop
F;88-771/Domain: myosin motor domain homology <MMOT>
F;178-185/Region: nucleotide-binding motif A (P-loop)
```

Query Match		6.2%; Score 321.5; DB 2; Length 1938;
Best Local Similarity		20.7%; Pred. No. 2.6e-06;
Matches		232; Conservative 184; Mismatches 438; Indels 269; Gaps 42;
Qy	73	QONVDVFAEDIHGTAERYAAARGVNI---HQOLLEHIRKLKPNQNTNPEGTSTGTPD 129
Db	851	EEENQAKEEMQKITERQQAETELKELOKHQTLAEEKTLLOEQLO-----AETELYA 904
Qy	130	EAAFLAERTPTDAESLLEKTPDEAARLVEGTSAKIQCLGKATSGKFEQ-----STE 180
Db	905	ESEEMRVLAACKQELBEILLHEMEARLEEDRQQL--QAERKMAQQMLDLEEQLEEE 962
Qy	181	ETPRKILRPTKETSEKFSWPAK-----ERSKLTWEEKETSVKTECVAGVTN 228
Db	963	EAARQKLQLEKVTAB-----AKIKLEDDILVMDQNSKLS---KERKLEERVSDITTN 1014
Qy	229	KTEVLEKGTSMIACPTKETSTKASTNV-----DVSSV-EPIFS 266
Db	1015	LAEBEEK-AKNLTCLKSHESMISELVRLKKEKSQELKRLKLEGDASDPHEQIAD 1073
Qy	267	LFGRTRTIENSQCTKVEEDFNATKIISKSAQNTVCLPDATYQKDITINHIEDQMPFS 326
Db	1074	LOAQIAELKMLAKKEELQAALARDDEIAQKNAL-----KKIRELEGHISDLQEDL 1127
Qy	327	ESK---REEDBEYSWDSGLSPESAKTQV--CIFESMYQKVMENRE----- 368
Db	1128	DSERAARNKAEKRDLDGEELE-ALKTELEDTLDSATQBELRAKRGQEVTVLKKALDEE 1186
Qy	369	-----VEELPEKPSAFKPAV-----EMQKTVPNKAPFELKNEOTLRAAQ 406
Db	1187	TRSHEAQVQEMRQKHQTQAVBELTQLEQFKKANKLDKSKOTLEKENADLAGELRVLQQA 1246
Qy	407	MPPSESKQKDDENSWSESPECVTQSKDVLPKATH--OKEFDTLSKLESVPKQGLL 464
Db	1247	KQEVHKKKLEVLQQLDQSKCSDGERARAEISDKVHKLQNEVSSTVGLMNEA----- 1299
Qy	465	KPTCGRKVSLPNKALEL-KDRETFKAESPDKDGLLKPTCGRKVSLPNKALKDRETLKA 523
Db	1300	-----EGKATKLVADVASIGSQDTQELLOEETROKLVNSTKLRQLEDERNSLQ 1349
Qy	524	ESPNDGLLKPTCGRKVSLPNKALELKDRETFKAAQMPFSESKQKDDENSWDFESFLET 583
Db	1350	DQDDEMEAKQNLERHVSITLNIQL-----SDSKK-----LQDFASTTIEV 1389
Qy	584	LLQNDVCLPKATHQKEFTLSKLEB-----SPDKDGLLKPTCGMKKISLPNKA----- 631
Db	1390	MEEG-----KKRLQKEMEGLSQYBEKAAAYDKLEKTKNRLQOELDDLVDLDNQRLVS 1444
Qy	632	-LELQDR-----ETPKAEVDSVSEFSLFGKPTTENSOSTKVEEDF 672
Db	1445	NLEKQKQKPDQLLABEKNISKYADRRDAEABAREKETALSARALE--EALAEKSEL 1502
Qy	673	NLTKE-----GATKVTVQOERDIGIIERAPQDQTNKMPSTSELGRKEDTKTSDSE 724
Db	1503	ERTNMLKAMEDLVSSKDDVGNVHLEKSKRALETQMEEMKTOLEESDDVQATEDAK 1562
Qy	725	I-ISVSDTONYECLPATYQKIKITNGKIESP---EKPSHFEPATEQNQSVPNKGLEW 780
Db	1563	LRLEV-----NMQAL-KGQFERDLOARDEQNEEKRRQLOQLH-EVETELEDERKQALAA 1616
Qy	781	KNKOTLRADSTTLKILDALPSCGERELKKNCEQITAKMEQMKNKFCVLQKELSEAKE 840
Db	1617	AKKKLEGLDKOLEAD--SAIKGRE--BAIKQLRKLQAMKD---FQRELDARA 1666
Qy	841	IKSQL-----ENQK--AKWEQELCSVRLPINEEERKRVNDILKEIRPE----- 883
Db	1667	SRDEIPATSKENEKAKSLADLMQLQEDLAAERAKQADLEKEBELABELASSLGRNT 1726
Qy	884	-EQRLKLELVK-HOLEQTFLRIQIDIELKSV-----TSNLQNVSHTHESENDLFHEN--- 931
Db	1727	LODEKRRLARIAQLBEEEEEEOGNMEASDRVRKATLQAEQSLNELATERSTAQNESA 1786

Qy	932	-----CMLKKEIAMLKLEVALTKHQHQVKENKYFEDIKLOE 968
Db	1787	ROOLERQNKELSKLOEVGAVKANKLSTVAALAKIAQLEQVEQEAEREKQAATKSLKQ 1846
Qy	969	KNAELQMTLKLKQKVTTKRASQYREQLKVLTAENTMLTSLKKE 1011
Db	1847	KDKKLKEVL-LQVEDERKMAEQYKEQAEGKNTKVQLKQLEEE 1888
RESULT 5		
JCS420		
smooth muscle myosin heavy chain 1 - mouse		
C:Species: Mus musculus (house mouse)		
C>Date: 10-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 09-Jul-2004		
C:Accession: JCS420		
R;Hasegawa, K.; Arakawa, E.; Oda, S.; Matsuda, Y.		
Biochem. Biophys. Res. Commun. 232, 313-316, 1997		
A:Title: Molecular cloning and expression of murine smooth muscle myosin heavy chains.		
A:Reference number: JCS420; MUID:97242182; PMID:9125171		
A:Accession: JCS420		
A:Molecule type: mRNA		
A:Residues: 1-1972 <HAS>		
A:Cross-references: UNIPROT:O08638; DDBJ:D85923; NID:g1945077; PIDN:BAA19690.1; PID:g194		
A:Experimental source: smooth muscle		
C:Comment: This protein plays a role in smooth muscle cell contraction.		
C:Superfamily: myosin heavy chain; myosin motor domain homology		
C:Keywords: nucleotide binding; P-loop		
F:88-771/Domain: myosin motor domain homology <MMOT>		
F:178-185/Region: nucleotide-binding motif A (P-loop)		
Query Match		6.2%; Score 321.5; DB 2; Length 1972;
Best Local Similarity		20.7%; Pred. No. 2.7e-06;
Matches		232; Conservative 184; Mismatches 438; Indels 269; Gaps 42;
Qy	73	QONVDVFAEDIHGTAERYAAARGVNI---HQOLLEHIRKLKPNQNTNPEGTSTGTPD 129
Db	851	EEENQAKEEMQKITERQQAETELKELOKHQTLAEEKTLLOEQLO-----AETELYA 904
Qy	130	EAAFLAERTPTDAESLLEKTPDEAARLVEGTSAKIQCLGKATSGKFEQ-----STE 180
Db	905	ESEEMRVLAACKQELBEILLHEMEARLEEDRQQL--QAERKMAQQMLDLEEQLEEE 962
Qy	181	ETPRKILRPTKETSEKFSWPAK-----ERSKLTWEEKETSVKTECVAGVTN 228
Db	963	EAARQKLQLEKVTAB-----AKIKLEDDILVMDQNSKLS---KERKLEERVSDITTN 1014
Qy	229	KTEVLEKGTSMIACPTKETSTKASTNV-----DVSSV-EPIFS 266
Db	1015	LAEBEEK-AKNLTCLKSHESMISELVRLKKEKSQELKRLKLEGDASDPHEQIAD 1073
Qy	267	LFGRTRTIENSQCTKVEEDFNATKIISKSAQNTVCLPDATYQKDITINHIEDQMPFS 326
Db	1074	LOAQIAELKMLAKKEELQAALARDDEIAQKNAL-----KKIRELEGHISDLQEDL 1127
Qy	327	ESK---REEDBEYSWDSGLSPESAKTQV--CIFESMYQKVMENRE----- 368
Db	1128	DSERAARNKAEKRDLDGEELE-ALKTELEDTLDSATQBELRAKRGQEVTVLKKALDEE 1186
Qy	369	-----VEELPEKPSAFKPAV-----EMQKTVPNKAPFELKNEOTLRAAQ 406
Db	1187	TRSHEAQVQEMRQKHQTQAVBELTQLEQFKKANKLDKSKOTLEKENADLAGELRVLQQA 1246
Qy	407	MPPSESKQKDDENSWSESPECVTQSKDVLPKATH--OKEFDTLSKLESVPKQGLL 464
Db	1247	KQEVHKKKLEVLQQLDQSKCSDGERARAEISDKVHKLQNEVSSTVGLMNEA----- 1299
Qy	465	KPTCGRKVSLPNKALEL-KDRETFKAESPDKDGLLKPTCGRKVSLPNKALKDRETLKA 523
Db	1300	-----EGKATKLVADVASIGSQDTQELLOEETROKLVNSTKLRQLEDERNSLQ 1349
Qy	524	ESPNDGLLKPTCGRKVSLPNKALELKDRETFKAAQMPFSESKQKDDENSWDFESFLET 583
Db	1350	DQDDEMEAKQNLERHVSITLNIQL-----SDSKK-----LQDFASTTIEV 1389
Qy	584	LLQNDVCLPKATHQKEFTLSKLEB-----SPDKDGLLKPTCGMKKISLPNKA----- 631
Db	1390	MEEG-----KKRLQKEMEGLSQYBEKAAAYDKLEKTKNRLQOELDDLVDLDNQRLVS 1444
Qy	632	-LELQDR-----ETPKAEVDSVSEFSLFGKPTTENSOSTKVEEDF 672
Db	1445	NLEKQKQKPDQLLABEKNISKYADRRDAEABAREKETALSARALE--EALAEKSEL 1502
Qy	673	NLTKE-----GATKVTVQOERDIGIIERAPQDQTNKMPSTSELGRKEDTKTSDSE 724
Db	1503	ERTNMLKAMEDLVSSKDDVGNVHLEKSKRALETQMEEMKTOLEESDDVQATEDAK 1562
Qy	725	I-ISVSDTONYECLPATYQKIKITNGKIESP---EKPSHFEPATEQNQSVPNKGLEW 780
Db	1563	LRLEV-----NMQAL-KGQFERDLOARDEQNEEKRRQLOQLH-EVETELEDERKQALAA 1616
Qy	781	KNKOTLRADSTTLKILDALPSCGERELKKNCEQITAKMEQMKNKFCVLQKELSEAKE 840
Db	1617	AKKKLEGLDKOLEAD--SAIKGRE--BAIKQLRKLQAMKD---FQRELDARA 1666
Qy	841	IKSQL-----ENQK--AKWEQELCSVRLPINEEERKRVNDILKEIRPE----- 883
Db	1667	SRDEIPATSKENEKAKSLADLMQLQEDLAAERAKQADLEKEBELABELASSLGRNT 1726
Qy	884	-EQRLKLELVK-HOLEQTFLRIQIDIELKSV-----TSNLQNVSHTHESENDLFHEN--- 931
Db	1727	LODEKRRLARIAQLBEEEEEEOGNMEASDRVRKATLQAEQSLNELATERSTAQNESA 1786

A26655	myosin heavy chain [similarity] - slime mold (Dictyostelium discoideum)	
N:Contains:	myosin Atpase [EC 3.6.4.1]	
C:Species:	Dictyostelium discoideum	
C>Date:	05-Oct-1988 #sequence revision 05-Oct-1988 #text_change 09-Jul-2004	
C:Accession:	A26655; A24728; S00250	
R:Warrick, H.M.; De Lozanne, A.;	Leinwand, L.A.; Spudich, J.A.	
Proc. Natl. Acad. Sci. U.S.A.	83, 9433-9437, 1986	
A:Title:	Conserved protein domains in a myosin heavy chain gene from Dictyostelium disco	
A:Reference number:	A26655; MUID:87092266; PMID:3540939	
A:Accession:	A26655	
A:Molecule type:	DNA	
A:Residues:	1-2116 <NAR>	
A:CROSS-references:	UNIPROT:P08799; GB:M14628; GB:M11938; NID:g167834; PIDN:AAA33227.1;	
R:DeLozanne, A.; Lewis, M.; Spudich, J.A.;	Leinwand, L.A.	
Proc. Natl. Acad. Sci. U.S.A.	82, 6807-6810, 1985	
A:Reference number:	A24728; MUID:86016788; PMID:3901008	
A:Accession:	A24728	
A:Molecule type:	DNA	
A:Residues:	2035-2116 	
R:Waggle, G.; Noegel, A.; Scheel, J.;	Gerisch, G.	
FEBS Lett.	227, 71-75, 1988	
A:Title:	Phosphorylation of threonine residues on cloned fragments of the Dictyostelium	
A:Reference number:	S00250; MUID:88112226; PMID:2828113	
A:Accession:	S00250	
A:Status:	nucleic acid sequence not shown	
A:Molecule type:	DNA	
A:Residues:	1734-1893 <WAG>	
C:Comment:	The rod domain is highly periodic, containing a pattern of 7-residue repeats	
C:Superfamily:	myosin heavy chain; myosin motor domain homology	
C:Keywords:	actin binding; ATP; coiled coil; hydrolase; nucleotide binding; P-loop; phosphate	
F:1-818/Domain:	globular head <HED>	
F:89-747/Domain:	myosin motor domain homology <MMOT>	
F:179-186/Region:	nucleotide-binding motif A (P-loop)	
F:819-2116/Domain:	alpha-helical rod <ROD>	
Query Match	6.2%; Score 318.5; DB 2; Length 2116;	
Best Local Similarity	20.4%; Pred. No. 3.9e-06;	
Matches 210; Conservative 181; Mismatches 376; Indels 263; Gaps 44;		
QY	73 QQNVDDFAEDIHGTAIRYAAARGVNYIHQQLLEHRL-----PKNPQNTNPEGTS--T 125	
DB	1201 KKKVELDLDKSAQLAEATAAQAQALDKKKLEQELSEVOTQLSEANNKVNDSNTKHL 1260	
QY	126 GTPDEAALPAERTPTDAESLLEKTPDEAARLVEGTSAKIQCLGKATSGKFGESTETPRK 185	
DB	1261 ETSFNNLKLEAEQAQAKALEK-----KRL--GLSELUKVN-----EQLEE---- 1302	
QY	186 ILRPTKETSEKFSWPAKERSRKITWBEKETSVKTECVAGVTENK--TEVLEKGTSNMIAC 243	
DB	1303 --KKQESNEK-----RKVDLEKEVSELKDQIEEBVASKKAVTEAKNKKSELDDEI 1351	
QY	244 PTKETSTKASTNVDSVSPFSLFTRTRIENSQCTKVEEDFNLATKIISKAAQNYTCL 303	
DB	1352 KRQYADVSSRD--KSVEQLTKLQAKN--EELNRTAEAEAGQLDRAERSKKA----- 1400	
QY	304 PDATYQDKITNHKIDQMFPESEKREDEEYSDGSLFESSAKTQVCIPESMYQKVM 363	
DB	1401 -EFDLEAEVKNLEETAKKVAEKAKKAAETDTRTSKSEL--DDAKN--VSEQYQVQIK 1454	
QY	364 EINREVEE----LPEKPSAFKPAVEMQKTPNKAFLKNEQTLRAAQMFPSSKQKQDE- 418	
DB	1455 RLNEELSELRSVLEADERCNSAIKAKTAESALESILKDE--IDAANNAKAAERKSKEL 1512	
QY	419 -----ENSWDES-----PQETVSQKDVLPKATHQKEFTLTSKLE---ESPVKDGLLK 465	
DB	1513 EVRVAELESLEDKSGTVNVEFRKKDA-----EIDDLRARLDRETESIKSD--- 1560	
QY	466 PTCGRKVSFLPNKALEKDRTEFKAESPDKDGLLKPTCGRKVSLPNKALEKDRTELKARS 525	
DB	1561 -----EDKKNTRKQFPADLEA-----KVSEAQRVWVTRDLRKKKLES 1596	
QY	526 PDND---GLLKPTCGR-KVSLPNKALE--LKDRTEFKAQMFPSSKQKQDDENSWDFES 579	

DB	1597 DIIDLSTQDTEKSRKIKKLEQTLAER---RAAEEGSSKAADDEEIRKQVWQEV 1653	
QY	580 FLETLLQNDVCLPKATHOK-----EPDTLSKLE-ESPDKDGLLKPTCGMKKISLPNKAL 632	
DB	1654 ELRAQLDSEERAAALNASEKKIKSLVAEDVEKQLEDEILAKDKLVKAKRALEVELEVRD 1713	
QY	633 ELKDRETFKAEE-----DVSSVESTFSLFGKPTTENSQSTKVEEDFNLTKEGATK 682	
DB	1714 QLEEEEDSRSELEDSKRLTTEVEDIKKYD-----AEVEQNTKLDE-----ARK 1758	
QY	683 TVTGOQERDIGIIRAPQDQTNMTSELGRK-----EDTKSTSDSEIISVSDTQNYEC 736	
DB	1759 KLTD-----DVTDLKKQLEDEKKKLAESERAKRLESENEDEFLAKLDAEVKNRSRAEK--- 1811	
QY	737 LPEATYQKEIKTTNGKIBESPEKSHFEPATEMQNSVPNKGLEWKNKQTLRADSTLSKI 796	
DB	1812 -DRKKYEKOLKQTKYKIND-----EAATYQTEIGAAKLE----- 1845	
QY	797 LDALPSCERBELKKNCEQITAKMQMKNKFCVLQKELSEAKEIKSLENOKAKWEQEL 856	
DB	1846 -----DQIDELSKLEQEQAKATQADK---SKTLEGEIDNLRQAQIEDS- 1886	
QY	857 CSVRLPLNQBEKRR---NVDILKEKIRPEEQLRKKLVKHOLEOTLRIDIELKSVTGN 913	
DB	1887 --GKIKVRLEKRALEGELEELRETVEAE-----DSKSEAEQSKRLVELELEEDARN 1938	
QY	914 LNQVSHTHESENDPLHENCMLKKEIA-----MLKLEVATLKHQ- 951	
DB	1939 LQKEIDAKEIAED---AKSNLQREIVAEKGRLEESIARTNSDRSRKRLAEAITALTAQV 1995	
QY	952 -----HQVKNKYFE-DIKILOKNAELQMTLKLKQKTYTKRASQYR-----EQ 994	
DB	1996 DAEQAKANQOIKENKIKETELKEYRKFGSEKTKTKEFLVWEKLETDYKRAKKEADEQ 2055	
QY	995 LKVLTAENTM 1004	
DB	2056 QORLTVENDL 2065	
RESULT 8		
A41604	myosin heavy chain, smooth muscle, long splice form - rabbit	
N:Contains:	myosin Atpase [EC 3.6.4.1]	
C:Species:	Oryctolagus cuniculus (domestic rabbit)	
C>Date:	31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 19-Apr-2002	
C:Accession:	A41604; A33501	
R:Babij, P.; Kelly, C.; Periasamy, M.		
Proc. Natl. Acad. Sci. U.S.A.	88, 10676-10680, 1991	
A:Title:	Characterization of a mammalian smooth muscle myosin heavy-chain gene: complete	
A:Reference number:	A41604; MUID:92073350; PMID:1961735	
A:Accession:	A41604	
A:Molecule type:	mRNA	
A:Residues:	1-1972 <BAB>	
A:CROSS-references:	GB:M77812	
R:Nagai, R.; Kuro-o, M.; Babij, P.;	Periasamy, M.	
J. Biol. Chem.	264, 9734-9737, 1989	
A:Title:	Identification of two types of smooth muscle myosin heavy chain isoforms by cdn	
A:Reference number:	A33501; MUID:89255535; PMID:2722872	
A:Accession:	A33501	
A:Molecule type:	mRNA	
A:Residues:	1455-1972 <NAG>	
A:CROSS-references:	GB:J04833; NID:g165519; PIDN:AAA31407.1; PID:g165520	
A:Experimental source:	smooth muscle	
A>Note:	examination by Southern blotting for the regions of difference between this isoform and	
	active splicing	
C:Superfamily:	myosin heavy chain; myosin motor domain homology	
C:Keywords:	actin binding; alternative splicing; ATP; coiled coil; hydrolase; methylated	
F:88-771/Domain:	myosin motor domain homology <MMOT>	
F:178-185/Region:	nucleotide-binding motif A (P-loop)	
F:559-572/Region:	actin binding #status predicted	
F:633-647/Region:	actin binding #status predicted	
F:844-1938/Domain:	coiled coil #status predicted <COI>	

R;Seelig, H.P.; Schranz, P.; Schroeter, H.; Wiemann, C.; Griffiths, G.; Renz, M.
Mol. Cell. Biol. 14, 2564-2576, 1994
A;Title: Molecular genetic analyses of a 376-kilodalton Golgi complex membrane protein
A;Reference number: A56539; MUID:94187728; PMID:7511208
A;Accession: A56539
A;Molecule type: mRNA
A;Residues: 1-3259 <SEE>
A;Cross-references: UNIPROT:Q14789; EMBL:X75304; NID:9405714; PIDN:CAA53052.1; PID:94057
C;Genetics:
A;Gene: GDB:GOLGB1; GCP; GCP371
A;Cross-references: GDB:454958
A;Map position: 3q13.31-3q13.31
C;Superfamily: ginsin
C;Keywords: coiled coil; Golgi apparatus; transmembrane protein
F:3238-3254/Domain: transmembrane #status predicted <TM>

Query Match 6.0%; Score 311; DB 1; Length 3259;
Best Local Similarity 19.3%; Pred. No. 1.4e-05;
Matches 235; Conservative 212; Mismatches 425; Indels 346; Gaps 48;

Qy 14 VQNKASLTPLLLAIQKRSKQTVFELLTKNANAFNE---SKTALMLAICGSGSEIVGM 70
Db 1101 LQDTNQIDLLQAEISENQAIQKLTSTNTDASDGSVALYKEITVWISPPCTGSGSEHWK 1160
Qy 71 LLOQNV---DVPAEDIHGITAERYAAAGVNYIHQLEHTR---KLPKNPONTNPEGTS 124
Db 1161 ELEBKILALEKEQLOKQLOQALTSKAILKAQAEKERHREELKQOKDYNRLQEQF- 1219
Qy 125 TGTDEAPLAERTPDPAESILELTPDEAARLVGTSKAIQCLGKATSGKFEQSTETPR 184
Db 1220 ---DEOSKENENTGDQLRQIQIQRRESIDGKLPSTQOQESC---SSTPGLEEPFKATEQ 1273
Qy 185 KILPRTKTSKFSWPAKERSKRTIWEKTSVKTECVAGVTPNKTVELEKGTG-NMIAC 243
Db 1274 HHTQPVLESNLCPDWPSHED-----ASALQGTSTVAQIKA 1309
Qy 244 PTKB-TSTKASTNVDSVSEVIFSLGTRTTENSQCTKVBDNFNLATIKISKAQNYTC 302
Db 1310 QKBEIAEKVLELVKAVST-----TSELTKKSEVFPLOEQINQKQGLE--- 1352
Qy 303 LPDATYQDKIATNHKIE-----DQMFPSKREE 332
Db 1353 -----TESLTKVSHAEVHAESIQKLESSQIQAGLEHRELQPKLDELQKLSKKE 1406
Qy 333 DEEYSWDSGSLFESSA-----KTQVCIPESMY-----QKWEINREVEEL 372
Db 1407 DVSYL---SGQLSEKAAALTKIQTEIIQEDDLIKALHTQLEMQAKEHDERIKQLQVLCM 1464
Qy 373 PEKP-----SAFKPAVE--MQKTPVKNKAFELKNEQTLR-----AAQMFPSSEKODDEE 419
Db 1465 KQKPEEIGESRAKQIQKQLQAALISRKALKENKSLQELSARGTIERLTKSLADVE 1524
Qy 420 NSWSESPCEVTSQKDYLPK-ATHQKFEFTLSKLESVPKDGILLKPTCGRKVSLPNKA 478
Db 1525 SQVSAQN-----KQKDTVLGALLQBERDKLITMDRSLLENQSSSC-----ES 1571
Qy 479 LELKDRETFKAESPKQGLLKPTCGRKVSLPNKALELKDR----- 518
Db 1572 LKL-----ALEGLTEDEKVLKVEISLSSKSTAEQKELQKVEYILLQSYENVSN 1627
Qy 519 -----ETLKASPDNDGLLKPTCGRK-----VSLPN 544
Db 1628 EAERIQHVAVRQEQELKYLKSTANKKETEKQLQEAQEQEEMKEKMKRFAKSKQ 1687
Qy 545 KALELKDRETFKAQMPFSESKQKDDENSWDFSFLETLLQNDVCLPKATH--QKFD 602
Db 1688 KILEBEENDRLRAEVHPAGDTAKE-----CMETLSSNASMKEELERVMKEYET 1737
Qy 603 LSGKLEB-SPDKGLLKPTCGMKISLPN-----KALELKDRETFKAED-----VS 646
Db 1738 LSKAFQSLMSKSLSEVQDLKHQI EDNVSKQANLEATEKHQNTVTEGTSQISPGET 1797
Qy 647 SVESFTSLFGKPT-TENSQSTK-----VEEDFNLTTK-EGATKTVTQOQRDGIIRAP 699

Db 1798 EQDLSMSRTPCTSESPSAKSNPAVSKDFSHDEINNNYLOQIDQLKRIAGLEE--- 1854
Qy 700 QDQNKMPTESELGKEDTK---STSDSEIISVSDTQNYECLPEATYQKEIKTTNGKTEE 755
Db 1855 EKQNKKEFSQTLNEKNTLLSQISTKQELKMLQEEVTKMLLNQIQEELSRTV-KUKE 1913
Qy 756 SPEKSPHFEPATEMONGVPNKGLEWKNKQTLRADSTTSLSKILDALPSCERGLKKNCE 815
Db 1914 TAE-----EEKDDLEERLMNQ-----LAEANGSI-----GNYCQ 1942
Qy 816 QIT---AKMEQMKKFCVLQKELSEAEIKSOLNQKAKWQELCSVRL-----PL 863
Db 1943 DVTDAQIKNELLESEMKKCVSELEEKQQLVKERTKVESEIRKEYLEYKIOGAKEPG 2002
Qy 864 NOBEKERNVDILKEKIRPEQRLKLVKHQ-----LEQTLRIQDI-----ELKSVT 911
Db 2003 NKSHAKELQ-ELLKEKQEVKQLOKQC-IRYQEKISALERTVKALEFVQTESQKDLKTK 2060
Qy 912 SNLNQ-VSHTHESENDL-----FHENCMLKKEI-----AMLKLEV 945
Db 2061 ENLAQAVEHRKKAQAEASFVLLDDTQSEARVLAADNLKLKKELOSNKESVKSQMKQD 2120
Qy 946 ATLKHQHQVKNKVFEDIKILQEK-----NAEQMTLKLKQKTVTKRASQY 991
Db 2121 EDLERRLEQAEKHLKQKQMKQEKLDALRRKRVHLEETIGETVTLAKKQKEV---QOL 2176
Qy 992 REQLKVLTAENTMLTSLK 1009
Db 2177 QENLDSTVTQLAAPTMSK 2194

RESULT 12
A61231
myosin heavy chain nonmuscle form A - human
N;Alternate names: cellular myosin heavy chain; myosin type 9; NMHC-A
N;Contents: myosin ATPase (BC 3.6.4.1)
C;Species: Homo sapiens (man)
C;Date: 12-May-1994 #sequence revision 14-Jul-1994 #text_change 09-Jul-2004
C;Accession: A61231; A34876; I52562; I61692
R;Simons, M.; Wang, M.; McBride, O.W.; Kawamoto, S.; Yamakawa, K.; Gdula, D.; Adelstein, C.
Circ. Res. 69, 530-539, 1991
A;Title: Human nonmuscle myosin heavy chains are encoded by two genes located on different chromosomes
A;Reference number: A61231; MUID:91316803; PMID:1860190
A;Accession: A61231
A;Molecule type: mRNA
A;Residues: 1-715 <STM>
A;Cross-references: UNIPROT:P35579; UNIPROT:Q9UMJ0; GB:M69180; NID:gl89029; PIDN:AAA6176
R;Saez, C.G.; Myers, J.C.; Shows, T.B.; Leinwand, L.A.
Proc. Natl. Acad. Sci. U.S.A. 87, 1164-1168, 1990
A;Title: Human nonmuscle myosin heavy chain mRNA: generation of diversity through alternative splicing
A;Reference number: A34876; MUID:90138958; PMID:1967836
A;Accession: A34876
A;Molecule type: mRNA
A;Residues: 715-1961 <SAE>
A;Cross-references: GB:M31013; NID:gl89035; PIDN:AAA36349.1; PID:gl89036
R;Toothaker, L.E.; Gonzalez, D.A.; Tung, N.; Lemons, R.S.; Le Beau, M.M.; Arnaout, M.A.; Blood 78, 1826-1833, 1991
A;Title: Cellular myosin heavy chain in human leukocytes: isolation of 5' cDNA clones, cDNA sequencing, and expression analysis
A;Reference number: I52562; MUID:92003925; PMID:1912569
A;Accession: I52562
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-52, 'EAT', 56-659, 'T', 661-868, 'T', 870-930, 'C', 932-1239, 'KG', 1242-1337 <RES>
A;Cross-references: GB:M81105; NID:gl89886; PIDN:AAA59888.1; PID:gl89886
R;Bement, W.M.; Hasson, T.; Wirth, J.A.; Cheney, R.E.; Mooseker, M.S.
Proc. Natl. Acad. Sci. U.S.A. 91, 6549-6553, 1994
A;Title: Identification and overlapping expression of multiple unconventional myosin genes
A;Reference number: A55758; MUID:94294418; PMID:8022818
A;Accession: I61692
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 182-218 <BEM>

1544 ELEDLQATADAKLREV---NLQAM-KAQFERDLQ---GRDEQSEKKKQLVLRQVREM 1599
QY 766 ATEMQNSVFNKGLEWKNKQTLRADSTTLTKILDA-----LPSCEGR 807
DB 1596 EAELEDERKQRSWAARAKKLENDLKLDEAHIDLSANKNRDEAIKQLRKLQAQMKDCWREL 1655
QY 808 ELKONCEOITA-----KMEQMKNFVQLKSELSEAKIISQLENQKAKWQELCSV- 859
DB 1656 DDTASREBILQAQENKELKSMEAEMIQLQBELAAARAKRQAQQRDELADETANSS 1715
QY 860 ---RLPLNQEERKRNVDILKEKI-RPEBOLRKLKLVKHQLEOTLRIDIELKSVTSNLN 915
DB 1716 GKALAL---EKKRR---LEARIAQLEEEBEEQNGTINLDRLLKXNLQIDQINADLN 1768
QY 916 -QVSHYTHESNDLFHENCMLKKEIAMLKLEVATLKHQHVKENKYFEDIKILQEKNAELQ 974
DB 1769 LERGHAKNEN---ARQQLERQNKELKVK- --LQEMEGTVSKYKASITALEAKIAQLE 1821
QY 975 MTL-----KLKQTV-----TKRASQYRQLKVLTAENTMLTSLK 1010
DB 1822 EQLDNETKERQAACKQVRRETEKKLQVLLQVDDERRNAEQYKQDAKSTLKLKQKOLE 1891
QY 1011 E 1011
DB 1882 E 1882

RESULT 13
I52300
giantin - human
N;Alternate names: gcp372
C;Species: Homo sapiens (man)
C;Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004
C;Accession: I52300
R;Schda, M.; Misumi, Y.; Fujiwara, T.; Nishioka, M.; Ikehara, Y.
Biochem. Biophys. Res. Commun. 205, 1399-1408, 1994
A;Title: Molecular cloning and sequence analysis of a human 372-kDa protein loc
A;Reference number: I52300; MUID:95100974; PMID:7802676
A;Accession: I52300
A;Status: preliminary; translated from GB/EMBL/DBD
A;Molecule type: mRNA
A;Residues: 1-3225 <RES>
A;Cross-references: UNIPROT:Q14789; GB:D25542; NID:g662389; PIDN:BAA05025.1; P
C;Superfamily: Giantin

Query Match 6.0%; Score 310; DB 2; Length 3225;
Best Local Similarity 19.3%; Pred. No. 1.5e-05;
Matches 235; Conservative 211; Mismatches 456; Indels 346; Gaps 48

QY 14 VQNKASITPLLAIQKSKQTVPELLTKNANAFNE---SKTALMLAICGSSEIVGM 70
DB 1067 LQDKTNQIDLLQAEISENQAIQKLIITSDASDGSVALYKETVWISPPCTGSEHWKP 1122
QY 71 LLOQNV---DVPAEDIHGITAERYAAAGVNYIHQOLLEHIR---KLPKNPQNTNPEGTS 124
DB 1127 ELEEKILALEKKEQLQKKQLEALTSKAILKKAQEKERHLREELKQKQDDYRNLQEQF- 1181
QY 125 TGTDEAAPLAERTPDTAESLLEKTPDDEARLVGTSAKIQCLGKATSGKGFQSTTEPTR 184
DB 1186 ---DEQKENENIGDQLRQLQIQVRESIDGKLPSDQOESC--SSTPLGEPLFLKATEQ 1233
QY 185 KILPTKETSFKFQWPAKERSRKITWEKETSVKTECVAGTVPNKTVELEKTS-NMLTAC 243
DB 1240 HHTQPVLESNLCPDWPVSHSD-----ASALQGTSTVAQIKA 1271
QY 244 PTKE-TSTKASTNVDVSSVEPISLFGTRTIENSQCTKVEBDFNLATKIISKSAQNYTC 302
DB 1276 QLKIEAKVELELKVSST-----TSELTKKSEVFQQLQKINQKGLE---- 1311
QY 303 LPDATYQKDKITINHKE-----DQMPFSSSKREE 332
DB 1319 -----IESLKTVSHEAEVHAESLQKLESSQLAGLEHLRELQPKLEOKLISKKE 137

```
Qy 333 DEEYSDSGSLPSSA-----KTQVCIPESMY-----QKVMINEIVEEL 372
Db 1373 DVSYL--SGQLSEKAALTKIQTEIEQEDLIKALHTQLEMQAKEHDEHRIKQLOVELCEM 1430
Qy 373 PEKP-----SAPPAVE--NOKTVPNKAFELKNEQTLR-----AAQMPFSESQKODDE 419
Db 1431 KQKEEIEEERAKQOIKRQLOAALISREKALKENKSQBELSLARGTIERLTLSLADVE 1490
Qy 420 NSWDESEPCETVSKQDVVLPK-ATHQKEFDTLSGKLESPVKDGLLKPTCGRKVSLPNKA 478
Db 1491 SQVSAQN-----KEKDTVLGALLQBERDKLITEMDRSLLENQSLSSC-----ES 1537
Qy 479 LELKDRFTKASPDQGLLKPTCGRKVSLPNKALELKDOR----- 518
Db 1538 LKL-----ALEGLTEDKELVKEIEBSLSSKSKIAESTEMOQEKHELOKEYEILLQSVNVS 1593
Qy 519 -----ETLKAESPNDGLLKPTCGRK-----VSLPN 544
Db 1594 EAEIRIQHVVEAVROEKQELYKLRSTEANKKETEKQLOEABOEMEEMKMKRKAQKQ 1653
Qy 545 KALELKDRFTKAAQMPFSESQKQDDEENSWDFESFLETLQNDVCLPKATH--QKEFDT 602
Db 1654 KILELEENDRLRAEVHPAGDTAKE-----CMETLSSNASMKEELERVQMEVET 1703
Qy 603 LSGKLEE--SPDKDGLLKPTCGMKISLPN-----KALELKDRFTKAEAD-----VS 646
Db 1704 LSKKFQSLMSKESLSIEVQDLKHQIEGNVSKQANLEATEKHNDQNTVTEBGTQSIPOET 1763
Qy 647 SVESFSLFGKPT--TENSQSTK-----VEEFPNLTK--EGATKTVTGOERDGIIEBP 699
Db 1764 EEOQLSGMSTRPTCSSEVPSAKSANPAVSKDFSSHDEINNYLOQIDQKERIAGLEE--- 1820
Qy 700 QDQTNKMPSTSELGRKEDTK-----STSDSEIISVSDTQNYECLPEATYQKEIKTTNGKIE 755
Db 1821 EKQKNKPSQTLNEKNTLLSQISTKQELQMLQOEVTKNLLNQIQEELSRVT--KLKE 1879
Qy 756 SPEKPSHPEPATEMNSVPNKGLBNKNQTLRADSTTSLKILDALPSCERGLKNONCE 815
Db 1880 TAE-----EEKDLEERLMNQ-----LAELNGSI-----GNVCQ 1908
Qy 816 QIT---AKMEQMKNFVLOKELSENAEIKSQLENQAKWQELCSVRL-----PL 863
Db 1909 DVTAQIKNELESEMKNLKKVCSELEBEKQOLVKEKTKVSEIRKYELEKIQGAQKEPG 1968
Qy 864 NOBEKERNVDILKEIRPEBQRLKLEVKHQ-----LEQTLRIODI-----ELKSVT 911
Db 1969 NKSHAKELQ-ELLKEKQEVQLOKDC--IRYQEKISALERTVKALEFPVQTESQKDLITK 2026
Qy 912 SNLQO-VSHTHSENDL-----PHENCMLKKEI-----AMLKLEV 945
Db 2027 ENLAQAVEHRKKAQAEALASFVLLDDDTQSEAAARVLADNLKLKXELQSNKESVKSMQKQD 2086
Qy 946 ATLKHQHOVKENKVFEDIKILQEK-----NAELMTLKLKQKTVTKRASQY 991
Db 2087 EDLERLEQAEBEHLKEKNQWQELDALREKRVHLEETIGTSIQVTLKNKQKEV-----QOL 2142
Qy 992 REQQLKVLTAENTMLTSKL 1009
Db 2143 QENLDSITVQLAAFTKSM 2160

RESULT 14
T14156
kinein-related protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T14156
R:Wood, K.W.; Sakowicz, R.; Goldstein, L.S.; Cleveland, D.W.
Cell 91, 357-366, 1997
A:Title: CENP-E is a plus end-directed kinetochore motor required for metaphase chromoso
A:Reference number: Z17893; MUID:98028574; PMID:9363944
A:Accession: T14156
A>Status: preliminary; translated from GB/EMBL/DBJ
```

A:Molecule type: mRNA
A:Residues: 1-2954 <WOO>
A:Cross-references: UNIPROT:O42263; EMBL:AF027728; NID:g2586070; PID:g2586071; PIDN:AAAC
C:Genetics:
A:Gene: XCENP-E
C:Superfamily: centromere protein E; kinesin motor domain homology

Query Match 5.9%; Score 307.5; DB 2; Length 2954;
Best Local Similarity 19.9%; Pred. No. 1.8e-05;
Matches 245; Conservative 200; Mismatches 402; Indels 383; Gaps 53;

```
Qy 17 KASLTPLLALAIQKSKQTV-----FLTYKANANAFNESCKTALMLAICE 62
Db 1211 KGHLTDSQLSIEKLEQLENLEVTETKLTQLQEBMKNTIERNELQTNFEDLK----- 1260
Qy 63 GSSIVGMLLQONVDVFAEDIHGITAERYAAAGVNYHQQLLEHRIK-----LP 112
Db 1261 AEHDSLKQDISENTEQSIIE-----TQDELRAAQBELREQKQLVDSFRQQLDSCVGSISP 1315
Qy 113 KNQNTNPEGTSTGTDPDEAAPLAERTPTDAESLLEKTPDEAARLVEGTSAKIQCLGKATS 172
Db 1316 NHDAVANQEKVSLG-----EVNSL-----QSEMLRGERDELQTSKALV 1354
Qy 173 GKFEQSTETPRKILRPTKETSEKPSWPAKERSRKITWEEKETSVKTECVAGVTPNKTEV 232
Db 1355 SELE-----LLRAHVKSVEGEN--LEITKLANGLEKEILGKSE-----ESEV 1394
Qy 233 LEKGTSNMIACTKETSTKASTNV--DVSSVEPIFSLFGTRTIENSQCTKVEEDFNLATKI 291
Db 1395 LKSMLENL-----KEDNNKLKQAEYSSENQFSL-----BEVFGSQKL 1435
Qy 292 ISK-----SAAQNTYTCLPDATY-----Q 309
Db 1436 VDEIEVLKAQKAAEERLEIKDRDYFELVQNTANLVGKLETPLOADHEEDSIDRRSEE 1495
Qy 310 KDIKTIHKKIEDQMPFSESKREDEEYSDSGSLPSSAKTQVCIPESMYQKVMINEIREV 369
Db 1496 MEIKVLGKLERNYLLERLQEEKLELS--NKLLEIKEMETSVLLKDDLOQKJESLSEN 1554
Qy 370 EELPEK--PSAPK---PAVEMQKTVPNKAPFLKNEQTLRAAQMPFSESQKQDDEENSWDS 424
Db 1555 IILKENIDTLKXHSDDTQAQLOKT--QOELQAKNLALAAASDNCPI-----TOEKETSADC 1608
Qy 425 ESPCETVSKQDVVLPKATHQK--EFDTL---SGKLESPVK-----DGLLKPTCGRKVSL 474
Db 1609 VHPLE---EKILLTTEELHQKTNEQKLEHKEKLEQAQVELKCEVEHLMKSMIESKSSL 1665
Qy 475 PNKALELKDRB-----TPKAESP 492
Db 1666 ESLOQEKHDTQOALLALKQOQVVTQBKKELQOTHEHLTAERVDHLKENIELGLNFKNEAQ 1725
Qy 493 DKDGLLKPTCGRKVSLPNKALE-----LKDR-----ETLKAESPNDGLL 532
Db 1726 QK--TTKEQC---LINENKELEQSHRLQCEIEBELMSLKDKKESALETLKSE----- 1773
Qy 533 KPTCGRKVSLPNKALE--LKDRFTKAAQ--MFPSESKQKDDDEENSWDFESFLET----- 583
Db 1774 -----QKVINLQEMWVMEELKNSQRTVIAERDQLOQDLRES--VENSIETQDDLRL 1826
Qy 584 -----LQNDVCLPKATHQKGFDTLSGKLESPDKDG 615
Db 1827 KAQBALQQQDKVQBELTSQISVLQEKISLLENQMLYNAVTVK---ETLSERDDLNQSKQH 1883
Qy 616 LLKPTCGMKISLPNKALELKDRFTFKA-----EDVSSVSESTF-----SLFG 656
Db 1884 LFSBIETLSLSLKEKSEFALQAEKDKDAARKTTDITEKISNIEBQLLQQAATNLKETEUYE 1943
Qy 657 KPT-----TENSQSTKVEEDFNLTTPKEGATKTVTQOQERD-----IGIER-- 697
Db 1944 RESLIQCKEQALANTEHLRETLKSKDLALGKWE-----QERDEAANKVIALTEKMS 1994
Qy 698 APODQTNKMPST---ELGRKE-----DTKSTSEIISVSDTQNYECLPEATYQKEIKT 748
```

Db 1995 SLEEQINENVTLKEGEGKETFYLRQPSKQOSSQMEELRESLTKYDLOLEAEKEISE 2054

Qy 749 TNGIERSPEKSPHFEPATEMNSVPNKGLEWKNKQTLRADSTYLSKILDALPSCERGR - 807

Db 2055 ATNEIKNLTKISSLEEEILQNASILNEAV--SERENLRHSKQOLVSELEQLSLTKSRD 2112

Qy 808 ----ELKDNCEQITAKNQWKNKFCVLQKELSEAKIQLQLENOKAKWQELCSVRLPL 863

Db 2113 HAPAQSKREKDEAVN-KIASIAEIKILTKEMDFRDSKESLQEQSHLSEELCTYKTEL 2171

Qy 864 NOEEKERNV-DILKEKIRPEQRLKKL-EVKHQLEQTLRIQDIELKSVTSNLMQVSHTH 921

Db 2172 QMLQKQKEDINNKLAKEKVEDELLOHLSLKEQLDQI-----QNELR-----NEKLRY 2221

Qy 922 ESENDLPHENCKMEIAMLKLEVATLKHQVQKKNYFEDIKILQENKABLQWTLKLKQ 981

Db 2222 E----LCEKMDIMEKEISVLR-----MONEPQOEEDVAERMDFILSRNOEI-----Q 2266

Qy 982 KTVTKRASQVREOLKVLTAENTMLTSKLKE 1011

Db 2267 ELMEXISAVYEQHTLJSS-----JSSSELQK 2292

RESULT 15

JC5837

C:Species: Rattus norvegicus (Norway rat)

C:Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004

C:Accession: JC5837

R:Toki, C.; Fujiwara, T.; Sohda, M.; Hong, H.S.; Misumi, Y.; Ikehara, Y.

Cell Struct. Funct. 22, 565-577, 1997

A:Title: Identification and characterization of rat 364-kDa Golgi-associated protein rec

A:Reference number: JC5837; PMID:98093490; PMID:9431462

A:Accession: JC5837

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-3187 <TOK>

A:Cross-references: UNIPROT:Q63714; DDBJ:D25543; NID:9516825; PID:BAA05026.1; PID:95168

C:Comment: This protein plays a role in the formation and maintenance of the characteris

C:Superfamily: giantin

F:49-549,624-1176,1238-1707,1763-3114/Domain: coiled-coil leucine zipper #status predic

F:3165-3187/Domain: membrane anchor #status predicted <MAD>

Query Match 5.8%; Score 302.5; DB 2; Length 3187;

Best Local Similarity 21.1%; Pred No. 3.2e-05;

Matches 253; Conservative 197; Mismatches 451; Indels 297; Gaps 54;

Qy 1 MVATLLSYGAVIEVQNK---ASLTPLLLAIQKRSKOTVEFLTKRNANANAFNESKCTAIM 57

Db 1453 LTKSLADVESQVSQVQNEKDALLGLKALLQERDKLIVE--MDKSLLENQSLGSCESLK 1510

Qy 58 LAICGSSSEIVGMLLQONVDVFAEDIHGITAERYAAAGVNYHQHLLHIRKLPKNQN 117

Db 1511 LAL--GG-----LTDEKELMKELSVRCISAEQWQEKHLEQYEVLLQSVEN 1561

Qy 118 TPEGFTGTGPDAAFLAERTPTDASLLEKTPDPAARLVGTSAKIQCLGKATSKFQ 177

Db 1562 VSN-----AERIQHVSVQEQEVYAKLSAESDKREKQLQDA--EQ 1606

Qy 178 STEETPRKILRPTKTSKFP-----SWPAYERSK 207

Db 1607 EMEEMKVRKFAKSKQKILELEENDRLRAEAPVGGANESMEALLSSNASLKEELER 1666

Qy 208 ITWEEKETSKTECVAGVTPNKEVLEKGTNN-----IACPTKETSTKASTNVDS 259

Db 1667 ITLEYTLKSEFALMA-----EKNTLSETRNLKQVEAQELQASLETTEKSDPKDV- 1721

Qy 260 SVPEIFSLFCTRIEN---SQCTKVEE-DFNLATKIISKSAQNYVCLPDA-TYQKDIKT 314

Db 1722 -IEEVTEAVVGKSEQEDSLSENAKLEDAEATLLANSAPCVSETFSHDDINNYLQQLDQ 1780

Qy 315 INHKIEDQMPSPSKSEKDEBEYSW-----DSGSLFPSSAK-TQVCIPESMYQKWEINRE 368

Db 1781 LKGRIAE-----LEMEKQKQDRELSQTLNEFKNALLTQISAKQSELKLEBEVAKINMLNQO 1836

Qy 369 VRELPEKPSAPKPAVE-----MQKTPVNPKAPEL-----KNEQ----- 400

Db 1837 IOEELSRVTKLETAEEDDLEERLMNQALBELNGSIYQYDVTDAQIKNQLESEMQN 1896

Qy 401 TLRAAQMFPESEKQKQDDEENSDSPCTVTSQKQVYLPK-----ATHQKEFDTL 450

Db 1897 LKRCVSELSEBKQOLVKEIKTVESEIRKE-----YMEKIQQAQKPGSKIHAKLEOEL 1949

Qy 451 SGKLESPPVKDGLLKPTC-----GRKVSIPN--KALELKDRETFFKAESPDKDGLLKPTCGR 504

Db 1950 ---LKEQEQEVKQLOQDCIRYLGRISALEKTVKALFVHTES-----QKD--LDATKG- 1997

Qy 505 KVSLPNKALELKDRETLKAESPNDGLLKPTCGRKVSLPNKALKALELKDRETFKAAQFPSE 564

Db 1998 ----NLAQAVEHHKQAQELSEFKILLDDTQSEAAARVLADNLKLK- KELSQNKESIKSQ 2051

Qy 565 SKQKDEENSDWPFLETLQNDVCLPKATHQKEFTDLSGKLEESPDKDGLLKPTCG- M 623

Db 2052 IKQKD-----EDLLRLLEQ-----ABEKIRKEKKNMQEKLDAHLHREKAHVEDTLAEI 2098

Qy 624 KISLPNKALELKDRETFKAEDVSSVESTSLFGKPTTENS-----QSTKVEEDF 672

Db 2099 QVSLTRKDKDMKELQ-----QSLDSTLQAALAAFTKSMSSLODDRDRVIDEAKKWEORF 2151

Qy 673 NLTTKEGATVTVGQOERDIGIIER---APQDOTNRMPTSELGRKEDTKTSDSEIISVS 729

Db 2152 G-----DAIQT-----KEEVRLKAEENCTALKDQLRQWTI---HMBELK-----ITVS 2191

Qy 730 ---DTQNYECLPEATVQKEIKITTKNGKIBESPEKPSHPPEPATEMNSVPNKGLEWKNK-Q 784

Db 2192 RLEHDKIEWESKQATELHQHQKAYDKLQEEENKELMSQLEAGQLYHDSKRNELTKLESELK 2251

Qy 785 TLRADSTTLKILDALPSCERGR--LK-----KDNCEQ-----ITAKME----- 822

Db 2252 SLKQDSTDLKNSLEKCKREHNNLEGIQKQEAQIONCKFCEQLETLDTLASRELTTLHD 2311

Qy 823 ---OMKNKFCVLQKELSE-----AKEIKSQLENQKAKWQELCSVRLPLNQE 866

Db 2312 EINVKEQKIISLLSGKEAIOVAIAELHQHSKEIK-ELENLLSQESEE-----NLTL 2364

Qy 867 BEKRNVDILKEKIRPEQRLKK-LEVKHQLEQTLR---IQDIELKSVTSNLMQVSHTH 921

Db 2365 EENKRAVEKTNQLTALETIKKESLEQKALDSPFKSMSSLOD--DRDRIVSDYRQLSERH 2423

Qy 922 -----ESENDLFHENCMLKKEIAML-----KLEVATLKHQHVKENKYFE 961

Db 2424 LSVILEKDELIDQAAAAENNKLEIRGLRGHMDLNSENAKLDAELIQYRRDLNEVITIK 2483

Qy 962 DIKILQENKABLQWTLKLKQKTV-----TKRASQVREQ-----LKVLTAEINTMLTSKLK 1010

Db 2484 DSQQRQLLEAQLQONKELRNECVKLEGRKGEAEKQSLQMSLDALQFENQGLSKK 2541

Search completed: February 19, 2005, 02:12:27

Job time : 38.5932 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 02:11:01 ; Search time 94.9578 Seconds
(without alignments)
3484.083 Million cell updates/sec

Title: US-09-602-362E-27

Perfect score: 5173

Sequence: 1 MVATLLSYGAVIEQVONKASL.....REQLKVLTAENTMLTSKLKE 1011

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5173	100.0	1011	16	US-10-408-765A-1557
2	5173	100.0	1239	13	US-10-007-805-557
3	5173	100.0	1239	14	US-10-076-622-577
4	5173	100.0	1239	14	US-10-124-805-577
5	5156	99.7	1225	14	US-10-177-293-332
6	3075	59.4	1341	13	US-10-007-805-565
7	3075	59.4	1341	14	US-10-076-622-565
8	3075	59.4	1341	14	US-10-177-293-334
9	3075	59.4	1341	14	US-10-124-805-565
10	3075	59.4	1341	15	US-10-058-270A-4
11	3069	59.3	1349	13	US-10-007-805-573
12	3069	59.3	1349	14	US-10-076-622-573
13	3069	59.3	1349	14	US-10-124-805-573

14	2374	45.9	1013	13	US-10-007-805-553	Sequence 553, App
15	2374	45.9	1013	14	US-10-076-622-553	Sequence 553, App
16	2374	45.9	1013	14	US-10-124-805-553	Sequence 553, App
17	2372.5	45.9	1095	9	US-09-834-759-493	Sequence 493, App
18	2372.5	45.9	1095	13	US-10-007-805-493	Sequence 493, App
19	2372.5	45.9	1095	14	US-10-076-622-493	Sequence 493, App
20	2372.5	45.9	1095	14	US-10-124-805-493	Sequence 493, App
21	2367	45.8	1002	9	US-09-604-287A-475	Sequence 475, App
22	2367	45.8	1002	9	US-09-834-759-475	Sequence 475, App
23	2367	45.8	1002	10	US-09-551-621-475	Sequence 475, App
24	2367	45.8	1002	13	US-10-007-805-475	Sequence 475, App
25	2367	45.8	1002	14	US-10-076-622-475	Sequence 475, App
26	2367	45.8	1002	14	US-10-124-805-475	Sequence 475, App
27	1758	34.0	661	13	US-10-007-805-552	Sequence 552, App
28	1758	34.0	661	14	US-10-076-622-552	Sequence 552, App
29	1758	34.0	661	14	US-10-124-805-552	Sequence 552, App
30	1756.5	34.0	743	9	US-09-834-759-494	Sequence 494, App
31	1756.5	34.0	743	13	US-10-007-805-494	Sequence 494, App
32	1756.5	34.0	743	14	US-10-076-622-494	Sequence 494, App
33	1756.5	34.0	743	14	US-10-124-805-494	Sequence 494, App
34	1751	33.8	650	9	US-09-825-301-25	Sequence 25, Appl
35	1751	33.8	650	9	US-09-604-287A-469	Sequence 469, App
36	1751	33.8	650	9	US-09-834-759-469	Sequence 469, App
37	1751	33.8	650	10	US-09-551-621-469	Sequence 469, App
38	1751	33.8	650	13	US-10-007-805-469	Sequence 469, App
39	1751	33.8	650	14	US-10-076-622-469	Sequence 469, App
40	1751	33.8	650	14	US-10-124-805-469	Sequence 469, App
41	1751	33.8	650	14	US-10-033-527-25	Sequence 25, Appl
42	1542	29.8	512	16	US-10-181-663-16	Sequence 16, Appl
43	1414.5	27.3	445	9	US-09-825-301-29	Sequence 29, Appl
44	1414.5	27.3	445	9	US-09-604-287A-473	Sequence 473, App
45	1414.5	27.3	445	9	US-09-834-759-473	Sequence 473, App

ALIGNMENTS

RESULT 1

US-10-408-765A-1557
; Sequence 1557, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1557
; LENGTH: 1011
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1557

Query Match 100.0%; Score 5173; DB 16; Length 1011;
Best Local Similarity 100.0%; Pred. No. 6.6e-263;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVATLLSYGAVIEQVONKASLTPLLLAIQKSKQVTFELLTKNANAFNESCCTALMLAI 60

Db 1 MVATLLSYGAVIEQVONKASLTPLLLAIQKSKQVTFELLTKNANAFNESCCTALMLAI 60

Qy 61 CEGSSEIVGMLLQNVDFVFAEDIHGITAERYAAARGVNIHQQLLEHRLKPKNPONTNP 120

Db 61 CEGSSEIVGMLLQNVDFVFAEDIHGITAERYAAARGVNIHQQLLEHRLKPKNPONTNP 120

Qy	121	EGTSTGTPDDEAAPIAERTPDPTAESLLEKTDPDEARLVEGTSAKIOCLGKATSGKFEOSTE	180
Db	121	EGTSTGTPDDEAAPIAERTPDPTAESLLEKTDPDEARLVEGTSAKIOCLGKATSGKFEOSTE	180
Qy	181	ETPRKILRPTKETSSEKSPAKERSRKITWEEKETSVKTECVAGVTNPKTEVLEKGTNNM	240
Db	181	ETPRKILRPTKETSSEKSPAKERSRKITWEEKETSVKTECVAGVTNPKTEVLEKGTNNM	240
Qy	241	IACPTKETSTKASTNDVSSVEPIFSLFGTRTIENSQCTKVEEDFNIAIKIISKAQNY	300
Db	241	IACPTKETSTKASTNDVSSVEPIFSLFGTRTIENSQCTKVEEDFNIAIKIISKAQNY	300
Qy	301	TCLPDATYQKDIKTINHIEQMPPESEKREDEEYSWDSGLPESAKTOVCIPESMYQ	360
Db	301	TCLPDATYQKDIKTINHIEQMPPESEKREDEEYSWDSGLPESAKTOVCIPESMYQ	360
Qy	361	KWMEINREVEELPEKPSAFKPAVEMQKTPNKAPELNEQTLRAAQMPPESEKQKDDEN	420
Db	361	KWMEINREVEELPEKPSAFKPAVEMQKTPNKAPELNEQTLRAAQMPPESEKQKDDEN	420
Qy	421	SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGLLKPTCGRKVSIPNKALE	480
Db	421	SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGLLKPTCGRKVSIPNKALE	480
Qy	481	LKDRETFKAESPDKDGLLKPTCGRKVSIPNKALELKDRETLKAESPNDGGLLKPTCGRKV	540
Db	481	LKDRETFKAESPDKDGLLKPTCGRKVSIPNKALELKDRETLKAESPNDGGLLKPTCGRKV	540
Qy	541	SLPNKALELKDRETFKAESPDKDGLLKPTCGRKVSIPNKALELKDRETLKAESPNDGGLLKPTCGRKV	600
Db	541	SLPNKALELKDRETFKAESPDKDGLLKPTCGRKVSIPNKALELKDRETLKAESPNDGGLLKPTCGRKV	600
Qy	601	DTLSGKLEESPDGGLLKPTCGRKVSIPNKALELKDRETLKAESPNDGGLLKPTCGRKV	660
Db	601	DTLSGKLEESPDGGLLKPTCGRKVSIPNKALELKDRETLKAESPNDGGLLKPTCGRKV	660
Qy	661	ENSQSTKVEEDFNLTKEGATKVTGQOERDGIITERAPOQOTNKMPSELGRKEDTKST	720
Db	661	ENSQSTKVEEDFNLTKEGATKVTGQOERDGIITERAPOQOTNKMPSELGRKEDTKST	720
Qy	721	SDSEIISVSDTQNYECLPEATYQKEITKTNGKIESPEKPSHFEPATEMONSVPNKGLEW	780
Db	721	SDSEIISVSDTQNYECLPEATYQKEITKTNGKIESPEKPSHFEPATEMONSVPNKGLEW	780
Qy	781	KNKQTLRADSTILSKILDALPSCERRELKKNCEQITAKMEOMKNKFCVLQKELSEAKE	840
Db	781	KNKQTLRADSTILSKILDALPSCERRELKKNCEQITAKMEOMKNKFCVLQKELSEAKE	840
Qy	841	IKSQLENQAKWEQELCSVRLPLNQEEKRRNVVDILKEKIRPEEQLRKKLEVKHOLEQTL	900
Db	841	IKSQLENQAKWEQELCSVRLPLNQEEKRRNVVDILKEKIRPEEQLRKKLEVKHOLEQTL	900
Qy	901	RIQDIELKSVTNLNOVSHSENDLPHENCMLKKEIAMLKLEVATLKHQOVKENYKF	960
Db	901	RIQDIELKSVTNLNOVSHSENDLPHENCMLKKEIAMLKLEVATLKHQOVKENYKF	960
Qy	961	EDIKILQENKAELOWTLKLKQKVTTKRASQYREQLKVLTAENTMLTSKLKE	1011
Db	961	EDIKILQENKAELOWTLKLKQKVTTKRASQYREQLKVLTAENTMLTSKLKE	1011

RESULT 2
US-10-007-805-577
; Sequence 577, Application US/10007805
; Publication No. US20020150581A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.
; APPLICANT: Panger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margarita
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C10
; CURRENT APPLICATION NUMBER: US/10/007,805
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 593
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 577
; LENGTH: 1239
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-805-577

Query Match	100.0%;	Score 5173;	DB 13;	Length 1239;
Best Local Similarity	100.0%;	Pred. No. 8.4e-263;		
Matches 1011;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MVATLSYGAVISVQNKASLTPLLLAIQKESKOTVEFLTKNANANAFNESCCTALMLAI	60	
Db	1	MVATLSYGAVISVQNKASLTPLLLAIQKESKOTVEFLTKNANANAFNESCCTALMLAI	60	
Qy	61	CEGSSEIVGMLLQONVDVFAEDIHGITAERYAAARGVNYIHQOLLEHRLPKNPQNTNP	120	
Db	61	CEGSSEIVGMLLQONVDVFAEDIHGITAERYAAARGVNYIHQOLLEHRLPKNPQNTNP	120	
Qy	121	EGTSTGTPDDEAAPIAERTPDPTAESLLEKTDPDEARLVEGTSAKIOCLGKATSGKFEOSTE	180	
Db	121	EGTSTGTPDDEAAPIAERTPDPTAESLLEKTDPDEARLVEGTSAKIOCLGKATSGKFEOSTE	180	
Qy	181	ETPRKILRPTKETSSEKSPAKERSRKITWEEKETSVKTECVAGVTNPKTEVLEKGTNNM	240	
Db	181	ETPRKILRPTKETSSEKSPAKERSRKITWEEKETSVKTECVAGVTNPKTEVLEKGTNNM	240	
Qy	241	IACPTKETSTKASTNDVSSVEPIFSLFGTRTIENSQCTKVEEDFNIAIKIISKAQNY	300	
Db	241	IACPTKETSTKASTNDVSSVEPIFSLFGTRTIENSQCTKVEEDFNIAIKIISKAQNY	300	
Qy	301	TCLPDATYQKDIKTINHIEQMPPESEKREDEEYSWDSGLPESAKTOVCIPESMYQ	360	
Db	301	TCLPDATYQKDIKTINHIEQMPPESEKREDEEYSWDSGLPESAKTOVCIPESMYQ	360	
Qy	361	KWMEINREVEELPEKPSAFKPAVEMQKTPNKAPELNEQTLRAAQMPPESEKQKDDEN	420	
Db	361	KWMEINREVEELPEKPSAFKPAVEMQKTPNKAPELNEQTLRAAQMPPESEKQKDDEN	420	
Qy	421	SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGLLKPTCGRKVSIPNKALE	480	
Db	421	SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGLLKPTCGRKVSIPNKALE	480	
Qy	481	LKDRETFKAESPDKDGLLKPTCGRKVSIPNKALELKDRETLKAESPNDGGLLKPTCGRKV	540	
Db	481	LKDRETFKAESPDKDGLLKPTCGRKVSIPNKALELKDRETLKAESPNDGGLLKPTCGRKV	540	
Qy	541	SLPNKALELKDRETFKAESPDKDGLLKPTCGRKVSIPNKALELKDRETLKAESPNDGGLLKPTCGRKV	600	
Db	541	SLPNKALELKDRETFKAESPDKDGLLKPTCGRKVSIPNKALELKDRETLKAESPNDGGLLKPTCGRKV	600	
Qy	601	DTLSGKLEESPDGGLLKPTCGRKVSIPNKALELKDRETLKAESPNDGGLLKPTCGRKV	660	
Db	601	DTLSGKLEESPDGGLLKPTCGRKVSIPNKALELKDRETLKAESPNDGGLLKPTCGRKV	660	
Qy	661	ENSQSTKVEEDFNLTKEGATKVTGQOERDGIITERAPOQOTNKMPSELGRKEDTKST	720	
Db	661	ENSQSTKVEEDFNLTKEGATKVTGQOERDGIITERAPOQOTNKMPSELGRKEDTKST	720	
Qy	721	SDSEIISVSDTQNYECLPEATYQKEITKTNGKIESPEKPSHFEPATEMONSVPNKGLEW	780	
Db	721	SDSEIISVSDTQNYECLPEATYQKEITKTNGKIESPEKPSHFEPATEMONSVPNKGLEW	780	


```
Qy 781 KKKOTLRADSTTLTKILDALPSCERGRRLKDNCEQITAKMEQMKNFVQLKELSEAKE 840
Db 781 KKKOTLRADSTTLTKILDALPSCERGRRLKDNCEQITAKMEQMKNFVQLKELSEAKE 840
Qy 841 KKSQLENQAKWEQELCSVRLPLNQEBEKKRRNVILKEKIRPEQOLRKLKLVKHOLEQTL 900
Db 841 KKSQLENQAKWEQELCSVRLPLNQEBEKKRRNVILKEKIRPEQOLRKLKLVKHOLEQTL 900
Qy 901 RIQDIELKSVTSNLNQVSHTHSENDLPHENCMLKKEIAMLKLEVALTLKHQHVKNKYF 960
Db 901 RIQDIELKSVTSNLNQVSHTHSENDLPHENCMLKKEIAMLKLEVALTLKHQHVKNKYF 960
Qy 961 EDIKILOEKNAELQMTLKLKQKTVTKRASQYREQLVLTAEENTMLTSKLKE 1011
Db 961 EDIKILOEKNAELQMTLKLKQKTVTKRASQYREQLVLTAEENTMLTSKLKE 1011

RESULT 3
US-10-076-622-577
; Sequence 577, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C11
; CURRENT APPLICATION NUMBER: US/10/076,622
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 577
; LENGTH: 1239
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-622-577

Query Match 100.0%; Score 5173; DB 14; Length 1239;
Best Local Similarity 100.0%; Pred. No. 8.4e-263;
Matches 1011; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVATLLSYGAVIEQVQNKASLTPLLLAIQKRSKQTVFLLTKNANANAFNESKCTALMLAI 60
Db 1 MVATLLSYGAVIEQVQNKASLTPLLLAIQKRSKQTVFLLTKNANANAFNESKCTALMLAI 60
Qy 61 CEGSSEIVGMLLQNVVDVFAEDIHGITAERYAAAGVNYIHQOLLEHIRKLPKPNQNTNP 120
Db 61 CEGSSEIVGMLLQNVVDVFAEDIHGITAERYAAAGVNYIHQOLLEHIRKLPKPNQNTNP 120
Qy 121 EGTSTGTPDEAAPLAERTPDTAEISLEKTPDEAARLVGTSKQIQLGKATSGKFEQSTE 180
Db 121 EGTSTGTPDEAAPLAERTPDTAEISLEKTPDEAARLVGTSKQIQLGKATSGKFEQSTE 180
Qy 181 ETPRKILRPTKETSFKSWPAKERSKITWBEKETSVKTECVAGVTPNKTVELEKGTSM 240
Db 181 ETPRKILRPTKETSFKSWPAKERSKITWBEKETSVKTECVAGVTPNKTVELEKGTSM 240
Qy 241 IACPTKETSSTKASTNVVDVSSVEPIPSLFGTRTIENSQCTKVEEDPNLATKIISKSAQNY 300
Db 241 IACPTKETSSTKASTNVVDVSSVEPIPSLFGTRTIENSQCTKVEEDPNLATKIISKSAQNY 300
Qy 301 TCLPDATYQDKIKTINHIEDQMFPSKREDEEYSWDSGLFESSAKTOVCIPESMYQ 360
Db 301 TCLPDATYQDKIKTINHIEDQMFPSKREDEEYSWDSGLFESSAKTOVCIPESMYQ 360
Qy 361 KWEINRREVEELPEKPSAFKPAVENQKTVPNKAFELKNEQTLRAAQMPFSSKQKDDEN 420
Db 361 KWEINRREVEELPEKPSAFKPAVENQKTVPNKAFELKNEQTLRAAQMPFSSKQKDDEN 420
Qy 421 SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGLLKPTCGRKVSLPNKALE 480
Db 421 SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGLLKPTCGRKVSLPNKALE 480
```

```
Db 421 SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGLLKPTCGRKVSLPNKALE 480
Qy 481 LKDRETFKAESPDKDGLLKPTCGRKVSLPNKALELKORETILKAESPNDGLLKPTCGRKV 540
Db 481 LKDRETFKAESPDKDGLLKPTCGRKVSLPNKALELKORETILKAESPNDGLLKPTCGRKV 540
Qy 541 SLPNKALELKDRETFKAQMPFSSKQKDDENSNDWDFESFLETLQNDVCLPKATHQKEF 600
Db 541 SLPNKALELKDRETFKAQMPFSSKQKDDENSNDWDFESFLETLQNDVCLPKATHQKEF 600
Qy 601 DTLGKLEESPDKGLLKPTCGMKISLPNKALELKDRETFKAEDVSSVESFSLFGKPTT 660
Db 601 DTLGKLEESPDKGLLKPTCGMKISLPNKALELKDRETFKAEDVSSVESFSLFGKPTT 660
Qy 661 ENSQSTKVEEDFNLTTTKEGATKTVTGOERDIGIIERAPODQTNKMPSELGRKEDTKST 720
Db 661 ENSQSTKVEEDFNLTTTKEGATKTVTGOERDIGIIERAPODQTNKMPSELGRKEDTKST 720
Qy 721 SDSIIISVSDTQNYECLPEATYQKEIKTTNGKIBESPEKPSHFEPATQMNSVFNKGLEW 780
Db 721 SDSIIISVSDTQNYECLPEATYQKEIKTTNGKIBESPEKPSHFEPATQMNSVFNKGLEW 780
Qy 781 KKKOTLRADSTTLTKILDALPSCERGRRLKDNCEQITAKMEQMKNFVQLKELSEAKE 840
Db 781 KKKOTLRADSTTLTKILDALPSCERGRRLKDNCEQITAKMEQMKNFVQLKELSEAKE 840
Qy 841 KKSQLENQAKWEQELCSVRLPLNQEBEKKRRNVILKEKIRPEQOLRKLKLVKHOLEQTL 900
Db 841 KKSQLENQAKWEQELCSVRLPLNQEBEKKRRNVILKEKIRPEQOLRKLKLVKHOLEQTL 900
Qy 901 RIQDIELKSVTSNLNQVSHTHSENDLPHENCMLKKEIAMLKLEVALTLKHQHVKNKYF 960
Db 901 RIQDIELKSVTSNLNQVSHTHSENDLPHENCMLKKEIAMLKLEVALTLKHQHVKNKYF 960
Qy 961 EDIKILOEKNAELQMTLKLKQKTVTKRASQYREQLVLTAEENTMLTSKLKE 1011
Db 961 EDIKILOEKNAELQMTLKLKQKTVTKRASQYREQLVLTAEENTMLTSKLKE 1011

RESULT 4
US-10-124-805-577
; Sequence 577, Application US/10124805
; Publication No. US20030166022A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C12
; CURRENT APPLICATION NUMBER: US/10/124,805
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 577
; LENGTH: 1239
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-124-805-577

Query Match 100.0%; Score 5173; DB 14; Length 1239;
Best Local Similarity 100.0%; Pred. No. 8.4e-263;
Matches 1011; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVATLLSYGAVIEQVQNKASLTPLLLAIQKRSKQTVFLLTKNANANAFNESKCTALMLAI 60
Db 1 MVATLLSYGAVIEQVQNKASLTPLLLAIQKRSKQTVFLLTKNANANAFNESKCTALMLAI 60
Qy 61 CEGSSEIVGMLLQNVVDVFAEDIHGITAERYAAAGVNYIHQOLLEHIRKLPKPNQNTNP 120
Db 61 CEGSSEIVGMLLQNVVDVFAEDIHGITAERYAAAGVNYIHQOLLEHIRKLPKPNQNTNP 120
Qy 121 EGTSTGTPDEAAPLAERTPDTAEISLEKTPDEAARLVGTSKQIQLGKATSGKFEQSTE 180
Db 121 EGTSTGTPDEAAPLAERTPDTAEISLEKTPDEAARLVGTSKQIQLGKATSGKFEQSTE 180
```

Db 121 EGTSTGTPDEAALPAAERTPDTAELSLEKTPDEAARLVGTSKATQCLGKATSGKFEQSTE 180
Qy 181 ETPRKILRPKTSEKTSWPAKERSRKITWEEKETSVKTECVAGVTNPKTEVLEKGTSM 240
Db 181 ETPRKILRPKTSEKTSWPAKERSRKITWEEKETSVKTECVAGVTNPKTEVLEKGTSM 240
Qy 241 IACPTKETSTKASTNDVSSVEPIPSLFGTTIENSQCTKVVEEDFNATKIISKSAQNY 300
Db 241 IACPTKETSTKASTNDVSSVEPIPSLFGTTIENSQCTKVVEEDFNATKIISKSAQNY 300
Qy 301 TCLPDATYQKDIKTINHKEIDQMPFSESREDEEYSWDSGLSPFESSAKTQVCIPESMYQ 360
Db 301 TCLPDATYQKDIKTINHKEIDQMPFSESREDEEYSWDSGLSPFESSAKTQVCIPESMYQ 360
Qy 361 KVMENINREVEELPEKPSAFKPAVEMQKTPVNPNAFELKNEQTLRAAQMPFSESREDEEYSWDSGLSPFESSAKTQVCIPESMYQ 420
Db 361 KVMENINREVEELPEKPSAFKPAVEMQKTPVNPNAFELKNEQTLRAAQMPFSESREDEEYSWDSGLSPFESSAKTQVCIPESMYQ 420
Qy 421 SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGKLLKPTCGRKVSPLNKALE 480
Db 421 SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGKLLKPTCGRKVSPLNKALE 480
Qy 481 LKDRETFKAESPDKDGLLKPTCGRKVSPLNKALELKDRETLKABSPDNDGLLKPTCGRKV 540
Db 481 LKDRETFKAESPDKDGLLKPTCGRKVSPLNKALELKDRETLKABSPDNDGLLKPTCGRKV 540
Qy 541 SLPNKALELKDRETFKAESPDKDGLLKPTCGRKVSPLNKALELKDRETLKABSPDNDGLLKPTCGRKV 600
Db 541 SLPNKALELKDRETFKAESPDKDGLLKPTCGRKVSPLNKALELKDRETLKABSPDNDGLLKPTCGRKV 600
Qy 601 DTLSGKLEESPVKDGKLLKPTCGRKVSPLNKALELKDRETFKAESPDKDGLLKPTCGRKV 660
Db 601 DTLSGKLEESPVKDGKLLKPTCGRKVSPLNKALELKDRETFKAESPDKDGLLKPTCGRKV 660
Qy 661 ENSQSTKVVEEDFNATKIISKSAQNY 720
Db 661 ENSQSTKVVEEDFNATKIISKSAQNY 720
Qy 721 SDSEIISVSDTQNYECLPEATYQKEIKTTNGKIESPEKPSFEPATQNTMPTSELGRKEDTKST 780
Db 721 SDSEIISVSDTQNYECLPEATYQKEIKTTNGKIESPEKPSFEPATQNTMPTSELGRKEDTKST 780
Qy 781 KXQTLRASTTLKSLKIDALPSCERGERELKDNCEQITAKMEQKMKFVLOKELSEAKE 840
Db 781 KXQTLRASTTLKSLKIDALPSCERGERELKDNCEQITAKMEQKMKFVLOKELSEAKE 840
Qy 841 KXQTLRASTTLKSLKIDALPSCERGERELKDNCEQITAKMEQKMKFVLOKELSEAKE 900
Db 841 KXQTLRASTTLKSLKIDALPSCERGERELKDNCEQITAKMEQKMKFVLOKELSEAKE 900
Qy 901 RIQDIELKSVTSLNQVSHTHSENDLFHNCMLKKEIAMLKLEVALTLKHQHVKNKYF 960
Db 901 RIQDIELKSVTSLNQVSHTHSENDLFHNCMLKKEIAMLKLEVALTLKHQHVKNKYF 960
Qy 961 EDIKILOEKNAELQMTLKLKQKVTTRASQYRQQLKVLTAENTMLTSKLKE 1011
Db 961 EDIKILOEKNAELQMTLKLKQKVTTRASQYRQQLKVLTAENTMLTSKLKE 1011

RESULT 5
US-10-177-293-332
; Sequence 332, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lililie, James
; APPLICANT: Lililie, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen

; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; FILE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 332
; LENGTH: 1225
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-332

Query Match 99.7%; Score 5156; DB 14; Length 1225;
Best Local Similarity 99.7%; Pred. No. 6.5e-262;
Matches 1008; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MVATLLSYGAVIEQVONKASLTPLLLAIQKRSKQTVFLLTKNANAFNFKSCTALMLAI 60
Db 154 MVATLLSYGAVIEQVONKASLTPLLLAIQKRSKQTVFLLTKNANAFNFKSCTALMLAI 213
Qy 61 CEGSEISVGMLLQONVDFPAEDIHGITAERYAAAGVNYHQQLLEHRLKPNQNTNP 120
Db 214 CEGSEISVGMLLQONVDFPAEDIHGITAERYAAAGVNYHQQLLEHRLKPNQNTNP 273
Qy 121 EGTSTGTPDEAALPAAERTPDTAELSLEKTPDEAARLVGTSKATQCLGKATSGKFEQSTE 180
Db 274 EGTSTGTPDEAALPAAERTPDTAELSLEKTPDEAARLVGTSKATQCLGKATSGKFEQSTE 333
Qy 181 ETPRKILRPKTSEKTSWPAKERSRKITWEEKETSVKTECVAGVTNPKTEVLEKGTSM 240
Db 334 ETPRKILRPKTSEKTSWPAKERSRKITWEEKETSVKTECVAGVTNPKTEVLEKGTSM 393
Qy 241 IACPTKETSTKASTNDVSSVEPIPSLFGTTIENSQCTKVVEEDFNATKIISKSAQNY 300
Db 394 IACPTKETSTKASTNDVSSVEPIPSLFGTTIENSQCTKVVEEDFNATKIISKSAQNY 453
Qy 301 TCLPDATYQKDIKTINHKEIDQMPFSESREDEEYSWDSGLSPFESSAKTQVCIPESMYQ 360
Db 454 TCLPDATYQKDIKTINHKEIDQMPFSESREDEEYSWDSGLSPFESSAKTQVCIPESMYQ 513
Qy 361 KVMENINREVEELPEKPSAFKPAVEMQKTPVNPNAFELKNEQTLRAAQMPFSESREDEEYSWDSGLSPFESSAKTQVCIPESMYQ 420
Db 514 KVMENINREVEELPEKPSAFKPAVEMQKTPVNPNAFELKNEQTLRAAQMPFSESREDEEYSWDSGLSPFESSAKTQVCIPESMYQ 573
Qy 421 SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGKLLKPTCGRKVSPLNKALE 480
Db 574 SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGKLLKPTCGRKVSPLNKALE 633
Qy 481 LKDRETFKAESPDKDGLLKPTCGRKVSPLNKALELKDRETLKABSPDNDGLLKPTCGRKV 540

Db 1112 LKVLIAENTMLTSKLE 1128

RESULT 7

US-10-076-622-565

Sequence 565, Application US/10076622

Publication No. US20030023036A1

GENERAL INFORMATION:

APPLICANT: Houghton, Raymond L.

APPLICANT: Sleath, Paul R.

APPLICANT: Persing, David H.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER

FILE REFERENCE: 210121.470C11

CURRENT APPLICATION NUMBER: US/10/076,622

CURRENT FILING DATE: 2002-02-13

NUMBER OF SEQ ID NOS: 627

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 565

LENGTH: 1341

TYPE: PRT

ORGANISM: Homo sapiens

US-10-076-622-565

Query Match 59.4%; Score 3075; DB 14; Length 1341;

Best Local Similarity 57.4%; Pred. No. 8.6e-153;

Matches 664; Conservative 76; Mismatches 145; Indels 272; Gaps 12;

QY 1 MVATLLSYGAVTEVQNKASLTPLLLAIQKRSQTVFILLTKNANANAFNSKCTALMLAI 60

Db 98 VVAKLLSHGAVLEVNKASLTPLLLSITKRSQIVFELLKNANANAVNKYCTALMLAV 157

QY 61 CGSSEIVGMLLQONVDVPAEDIGHITAIRYAAAGVNYIHQOLLEHRIKLPKPNQNTNP 120

Db 158 CHGSSEIVGMLLQONVDVPAEDICGVTAHYAVTCGFHHIHEQIMEYIRKLSKHQNTNP 217

QY 121 EGTSTGTPEEAAPLA-----ERTPDTABSLLEKTPDEAARLVE 158

Db 218 EGTSGTPEEAAPLAERTPDTABSLVEKTPDEAAPLVERTPDTABSLVEKTPDEAASLVE 277

QY 159 GTSAKIQCLGKATSGKFEQSTETPRKILRPKTSEKFSWP----- 200

Db 278 GTSDKIQLCEKATSGKFEQSAETPREITSPAKETSEKFTWPAKGRPKRIAWKEDTPR 337

QY 201 -----AKRSRKITWEEKETSVKTECVAGVTPNKTVELEKGTSMIACP 244

Db 338 EIMSPAKETSEKFTWAAKGRPKRIAWKETTPTKTCVAVTSNKTVKLEGRSKMIACP 397

QY 245 TKETSTKASTNVDVSSVEPIFSLFGTRTIENSQCTKVEBDFNLATKIISKSAAYNTCLP 304

Db 398 TRSSSTKASAN----- 408

QY 305 DATYQDKITINHKBQDMPFSSKEBDEYVSWGSLFESSAKTQVCIPESMTQKWE 364

Db 409 -----DQRPSSKEBDEYVSCDSRSLFESSAKTQVCIPESITQKWE 452

QY 365 INREVELPKPSAFPAVEMQTVNKAFELKNEOTLRAAQWPFSSKQKDEENSWDS 424

Db 453 INREVEPPKPSAFPAEMQNSVFNKAFELKNEOTLRAQWPFPPESKQKDEENSWDS 512

QY 425 EPCETVSKQDVLPKATHQKEDTLISGLKEESPVRDGLLKPTCGRKVSLPNKALELKDOR 484

Db 513 ESI CETVSKQDVLPKATHQKEDTLISGLKEESPVRDGLLKPTCGRKVSLPNKALELKDOR 572

QY 485 ETPKAESPDKGLLKPTCGRKVSLPNKALELKDRETILKAESEPDNDGLLKPTCGRKVSLPN 544

Db 573 QTFKAEPKPSAFEP-----ATEMQK-----SVFN 598

QY 545 KALELKDRETILKAESEPDNDGLLKPTCGRKVSLPNKALELKDRETILKAESEPDNDGLLKPT 604

Db 599 KALELKDRETILKAESEPDNDGLLKPTCGRKVSLPNKALELKDRETILKAESEPDNDGLLKPT 658

QY 605 GKLEESPDKGLLKPTCGMKISLPNKALELKDRETILKAESEPDNDGLLKPTCGRKVSLPN 664

Db 659 GKLESPVRDGLLKANCGMKVSIPTKALELMDQWTFKAEPPEK-PSAF-----EPAIEMQK 713

QY 665 STKVEEDFNLTTEGATKTVTCQOERDGIIRAPQDQTNKMPSTSELGRKEDTKSTDSGE 724

Db 714 SV-----PNKALELKNQOTLRADILP-----SESKQKQYEESWDS 751

QY 725 IISVSDTONVECLPEATYQKEIKTTNGKIEES----- 756

Db 752 SLCETVSKQDVCLPKATHQKEDIDKINGKLEESPNDGFLKAPCRMKVSIPTKALELMDQW 811

QY 757 -----PEKPSHPPEPATMONGVNPKNGLKWNKQTLRA----- 788

Db 812 TPKAEPPEKPSAFPAIEMQKSVFNKALELKNQOTLRADQMPFSSKQKKEENSWDS 871

QY 789 -----DSTTLTKILDALPSCERGLKKNQCNQITA 819

Db 872 LRETVSQKDVCPKATHQKEMDKISKLEEDSTLSKILDTVHSCERARLQKDHCEQRTG 931

QY 820 KMEQMKKFCVLQKELSEAKIKSQLENQKAWQELCSVRLPLNQEEKRRNVLDILKSK 879

Db 932 KMEQMKKFCVLKLLSEAKIKSQLENQKAWQELCSVRLTLNQEEKRRNADILNEK 991

QY 880 IRP-----EQLRKKLEVKHOLEQTLRIQDIELKSVTSLNQVSHTHESENDLFHENCML 934

Db 992 IREELGRIERQHRKELEVKQLEQALRIQDIELKSVESNLNQVSHTHESENDLFHENCML 1051

QY 935 KKEIAMLKLEVATLKHOHQVKNKYFEDIKLOEKVLAELQMTLKKOKTVTKRASQVREQ 994

Db 1052 KKSIAMLKLSIATLKHOQYQENKYNFEDIKLEKVAELQMTLKKKEAELQMTLKKESILTKEASQVSGQ 1111

QY 995 LKVLIAENTMLTSKLE 1011

Db 1112 LKVLIAENTMLTSKLE 1128

RESULT 8

US-10-177-293-334

Sequence 334, Application US/10177293

Publication No. US20030124128A1

GENERAL INFORMATION:

APPLICANT: Lillie, James

APPLICANT: Glatt, Karen

APPLICANT: Zhao, Xumei

APPLICANT: Gannavarpu, Manjula

APPLICANT: Kamatkar, Shubhangi

APPLICANT: Mertens, Maureen

APPLICANT: Myer, Vic

APPLICANT: Wang, Youzhen

APPLICANT: Xu, Yongyao

APPLICANT: Hoersch, Sebastian

APPLICANT: Monahan, John

APPLICANT: Meyers, Rachel E.

APPLICANT: Bast Jr., Robert C.

APPLICANT: Hortobagyi, Gabriel N.

APPLICANT: Pusztai, Lajos

APPLICANT: Sahin, Aysegul

APPLICANT: Mills, Gordon B.

TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,

TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER

FILE REFERENCE: MRI-038

CURRENT APPLICATION NUMBER: US/10/177,293

CURRENT FILING DATE: 2002-06-21

PRIOR APPLICATION NUMBER: US 60/299,887

PRIOR FILING DATE: 2001-06-21

PRIOR APPLICATION NUMBER: US 60/301,572

PRIOR FILING DATE: 2001-06-27

PRIOR APPLICATION NUMBER: US 60/306,501

PRIOR FILING DATE: 2001-07-18

PRIOR APPLICATION NUMBER: US 60/325,002

PRIOR FILING DATE: 2001-09-25

PRIOR APPLICATION NUMBER: US 60/362,585

Db 409 -----DORFSESKQEBDEYSCDSLSLFESSAKIQVCIPESIQKVM 452
Qy 365 INREVEELPEKPSAFKPAVENQKTPVNPKAFAELKNEQTLRAAQMPSPSSKQKDDSENSWDS 424
Db 453 INREVEEPKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPFPPSPSSKQKDYENSWS 512
Qy 425 ESPCETVSQKDVYLPKATHQKFEFTLSGKLEESPVKGLLKPCTGCRKVSIPNKALELKDR 484
Db 513 ESLCETVSQKDVCLPKATHQKEIDKINGKLEESPNKGLLKATCGMKVSIPTKALELKM 572
Qy 485 ETKAESPDGKGLLKPTGCRKVSIPNKALELKDRTELKAEPSDNDGLLKPTGCRKVSIPN 544
Db 573 QTFKAEPKPSAFEP-----ATENQK-----SVPN 598
Qy 545 KALELKDRETFKAQMPSPSSKQKDDSENSWDSFLETLQNDVCLPKATHQKEFTLS 604
Db 599 KALELKNEQTLRADDEILPSPSSKQKDYENSWSDESCLCETVSQKDVCLPKAAHQKEIDKIN 658
Qy 605 GKLEESPDGKGLLKPTCGMKISLPNKALELKDRETFKAEDVSSVESTFSFGKPTTENSQ 664
Db 659 GKLESPVKDGLLKANCCKMKSIPNKALELMDMTFKAEPPEK-PSAF-----EPAIEMQK 713
Qy 665 STKVEEDFNLTTKGATKVTGQOERDGIITERAFQDQTNKMPSELGRKEDTKSTSDSE 724
Db 714 SV-----PNKALELKNEQTLRADEILP-----SESKQKDYESSWDS 751
Qy 725 IISVSDTQNYECLPATYQKIKITNGKIBES----- 756
Db 752 SLCETVSQKDVCLPKATHQKEIDKINGKLEESPNNDGFLKAPCRMKVSIPTKALELMDQ 811
Qy 757 -----PEKPSHFEATQMNSVNPNGLEWKNKOTLRA----- 788
Db 812 TFKAEPPKPSAFEPKAIEMQNSVPNKALELKNEQTLRADQMPSPSSKQKDYENSWSDES 871
Qy 789 -----DSTLLSKILDALPSCERGERELKDKCEQITA 819
Db 872 LBRETVSQKDVCPKATHQKEMDKISGKLESDSTLSKILDTVHSCERARELQKDHCEQRTG 931
Qy 820 KNEQMKNFVLOKELSEBAKISOLENQAKWEQELCSVRPLNQBEKERNVDILKEK 879
Db 932 KNEQMKNFVLOKELSEBAKISOLENQAKWEQELCSVRPLNQBEKERNVDILKEK 991
Qy 880 IRP-----EQLRKLKLVKHOEQLTRIQDIELKSVTSNLAQVSHTHESNDLPHENCM 934
Db 992 IREELGRIEBOHRKEVQKLEQALRIQDIELKSVTSNLAQVSHTHENYLLHENCM 1051
Qy 935 KKEIAMLKLEVATLKHQHVKENKVFEDIKILQEKNAELQMTLKLKQKVTTKRASQYREQ 994
Db 1052 KKEIAMLKLEIATLKHQYQEKENKVFEDIKILKKNELQMTLKLKESLTKRASQYSGQ 1111
Qy 995 LKVLTAENTMLTSKLKE 1011
Db 1112 LKVLTAENTMLTSKLKE 1128

RESULT 10
US-10-058-270A-4
; Sequence 4, Application US/10058270A
; Publication No. US20040029114A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Afari, Daniel
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Breast Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Breast Cancer
; FILE REFERENCE: 018501-005210US
; CURRENT APPLICATION NUMBER: US/10/058,270A
; CURRENT FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: US 60/263,965
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: US 60/265,928

; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: US 09/829,472
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 60/282,698
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 60/288,590
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,443
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1341
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-058-270A-4

Query Match 59.4%; Score 3075; DB 15; Length 1341;
Best Local Similarity 57.4%; Pred. No. 8.6e-153;
Matches 664; Conservative 76; Mismatches 145; Indels 272; Gaps 12;

Qy 1 MVATILSYGAVIEVQNKASLTPLLLAIOKESKOTVEFLTKNANAFNESCCTALMLAI 60
Db 98 VVAKLLSHGAVIEVHNKASLTPLLLSITKSEQIVFLLIKNANANAVNKYKCTALMLAV 157
Qy 61 CEGSSEIVGMLQONVDVFAEDIHGITAERYAAAGVNYIHOQLLEHRLKLPKNQNTNP 120
Db 158 CHGSSEIVGMLQONVDVFAADICGVTAEHYAVTCGFPHIHEQIMEYIRKLSKNHQNTNP 217
Qy 121 EGTSTGTPDRAAPLA-----ERTPDAESLLEKTPDEAARLVE 158
Db 218 EGTSGATPDRAAPLAERTPDTAESLVEKTPDEAAPLVERTPDTAESLVEKTPDEAASLVE 277
Qy 159 GTSNKOICLCKATSGKPEQSTESTPRKILRPTKETSKEFSP----- 200
Db 278 GTSKIOCLEKATSGKPEQSAETPREITSKETSKEFTWPAKGRPRKIAWEKEDTPR 337
Qy 201 -----AKERSRKITWEKETSVKTECVAGVTPNKTVELEKGTSNMIACP 244
Db 338 EIMSPAKETSEKTPWAKGRPRKIAWEKETPVKTGCVARVTSNKTVELEKGRSKMIACP 397
Qy 245 TKETSTKASNVDSVVEPIPSLPGTFTIENSQCTKVEEDFNLAATKIISSAAQNYTCLP 304
Db 398 TKESSTKASAN----- 409
Qy 305 DATYQKDIKTINHKEIDQMPSPSSKQREDEYSGWDSGLSFESSAKTOVCIPESIQKVM 364
Db 409 -----DQRPSPSSKQREDEYSCDSLSLFESSAKIQVCIPESIQKVM 452
Qy 365 INREVEELPEKPSAFKPAVENQKTPVNPKAFAELKNEQTLRAAQMPSPSSKQKDDSENSWDS 424
Db 453 INREVEEPKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPFPPSPSSKQKDYENSWS 512
Qy 425 ESPCETVSQKDVYLPKATHQKFEFTLSGKLEESPVKGLLKPCTGCRKVSIPNKALELKDR 484
Db 513 ESLCETVSQKDVCLPKATHQKEIDKINGKLEESPNKGLLKATCGMKVSIPTKALELKM 572
Qy 485 ETKAESPDGKGLLKPTGCRKVSIPNKALELKDRTELKAEPSDNDGLLKPTGCRKVSIPN 544
Db 573 QTFKAEPKPSAFEP-----ATENQK-----SVPN 598
Qy 545 KALELKDRETFKAQMPSPSSKQKDDSENSWDSFLETLQNDVCLPKATHQKEFTLS 604
Db 599 KALELKNEQTLRADDEILPSPSSKQKDYENSWSDESCLCETVSQKDVCLPKAAHQKEIDKIN 658
Qy 605 GKLEESPDGKGLLKPTCGMKISLPNKALELKDRETFKAEDVSSVESTFSFGKPTTENSQ 664
Db 659 GKLESPVKDGLLKANCCKMKSIPNKALELMDMTFKAEPPEK-PSAF-----EPAIEMQK 713
Qy 665 STKVEEDFNLTTKGATKVTGQOERDGIITERAFQDQTNKMPSELGRKEDTKSTSDSE 724
Db 714 SV-----PNKALELKNEQTLRADEILP-----SESKQKDYESSWDS 751

```
Qy 725 IISVSDTONYECLEPATYQKEIKTTNGKIBES----- 756
Db 752 SLCEVTSQKQVCLPKATHQKEIDKINGKLEESPNDGFLKAPCRMKVSIPTKALELMDMQ 811
Qy 757 -----PEKPSHFEPATQMNSVFNKGLWKNKOTLRA----- 788
Db 812 TFKAEPPEKPSAFPAIEMQSVNKALELQNEQTLRADQMPFSESQKKVENSWSSES 871
Qy 789 -----DSTTLKILDALPSCGERELKKNCEQITA 819
Db 872 LRETVSQKQVCPKATHQKEMDKISGKLEDSTLSKILDTVHSCERARELQKHCEQRTG 931
Qy 820 KMEQMKKFCVLOKELSEAKIKSQLENOKAKWEQELCSVRLPLNOBEERKNVDILKEK 879
Db 932 KMEQMKKFCVLLKKLSAKIKSQLENQKVKEQELCSVRLTLNQBEERKNADILNEK 991
Qy 880 IRP-----EEQRLKLEVKHOLEQTLRIQDIELKSVTSNNOVSHTHSENDLFHENCML 934
Db 992 IRBELGRIEEQRHRELEVKQLEQALRIQDIELKSVTSNNOVSHTHSENDLFHENCML 1051
Qy 935 KKEIAMLKLEVATLKHOQVKNKYFEDIKILQEKNAELQMTLKLKQKVTVKRASQYREQ 994
Db 1052 KKEIAMLKLEIATLKHOQVKNKYFEDIKILKEKNAELQMTLKLKESLTKRASQYSGQ 1111
Qy 995 LKVLTAENTMLTSKLKE 1011
Db 1112 LKVLIAENTMLTSKLKE 1128

RESULT 11
US-10-007-805-573
; Sequence 573, Application US/10007805
; Publication No. US20020150581A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margaita
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; CURRENT APPLICATION NUMBER: US/10/007,805
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 593
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 573
; LENGTH: 1349
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-805-573

Query Match 59,3%; Score 3069; DB 13; Length 1349;
Best Local Similarity 57,3%; Pred. No. 1.8e-152;
Matches 663; Conservative 76; Mismatches 146; Indels 272; Gaps 12;

Qy 1 MVATLLSYGAVIEQNKASLTPLLLAIQKRSQTVFELLTKNANANAFNSKCTALMLAI 60
Db 106 VVAKLLSHGAVIEVHNKASLTPLLLSITKRSQIVFELLTKNANANAVNKCTALMLAV 165
Qy 61 CEGSSEIVGMLLQNVDFABDIHGITAERYAAAGVNYIHQQLLEHRLPKPNQNTNP 120
Db 166 CHGLSEIVGMLLQNVDFVFAADIQGVTAERYAVTCGFHIEHQIMEVIRKLSKNHNTNP 225
Qy 121 EGTSTGTPDEAFLA-----ERTPDTAESLLEKTPDPAALVE 158
Db 226 EGTSGTGPDEAFLAERTPDTAESLVEKTPDPAALVETPDTAESLVEKTPDPAALVE 285
```

RESULT 12
US-10-076-622-573
; Sequence 573, Application US/10076622

```
Qy 159 GTSAKIOCLGKATSGKFEQSTETPRKILRPTKETSSEKFSWP----- 200
Db 286 GTSDKIOCLEKATSGKFEQSAEETPREITSPAKETSEKFTWPAKGRPRKIAWEKKEDTPR 345
Qy 201 -----AKERSKITWEEKETSVKTECVAGVTPNKTEVLEKGTSMIACP 244
Db 346 EIMSPAKETSEKFTWAAKGRPRKIAWEKKETPVKTCVAVRTSNKTKVLEKGRSMIACP 405
Qy 245 TKETSTKASTNVDSVSSVEPIPSLFGTRTIENSQCTKVEEDFNLATKIISKAAQNYTCLP 304
Db 406 TKESSTKASAN----- 416
Qy 305 DATYQDKITINHIEDQMPFSESREDEYSWDSGLPFESSAKTQVCIPESMYQKWE 364
Db 417 -----DQRFPSSEKQEBEDEYSCDSRSLPSSAKIOVCIPEIYQKWE 460
Qy 365 INREVELPEKPSAFKPAVEMQKTPVNPKAPKELKNEQTLRAAQMPFSESQKODDENSWDS 424
Db 461 INREVEPPPKPSAFKPAIEMQSVNKALELQNEQTLRADQMPFSESQKDYBENSWS 520
Qy 425 ESPCETVSQKDVYLPKATHQKEFDTLGKLEESPVKDGLLKPTCGRKVSLPNKALELKDOR 484
Db 521 ESLCETVSQKDVCLPKATHQKEIDKINGKLEESPNKDGLLKATCGMKVSIPTKALELKDOR 580
Qy 485 ETFAESPDKDGLLKPTCGRKVSLPNKALELKDRETLKAESPNDGLLKPTCGRKVSLPN 544
Db 581 QTFKAEPGKPSAPEP-----ATEMQK-----SVPN 606
Qy 545 KALELKDRETFKAQMPFSESQKQDDENSWDFSFLETLLONDVCLPKATHQKEFDTL 604
Db 607 KALELKNEQTLRADQILPSESQKQDYBENSWDTSLSCTVTSQKDVCLPKAAHQKEIDKIN 666
Qy 605 GKLEESPDKDGLLKPTCGMKISLPNKALELKDRETFKAEDVSSVESFSLFGKPTTNSQ 664
Db 667 GKLEESPVKDGLLKANCGMKVSIPTKALELMDQMTFKAEPPEK-PSAF-----EPAIEMQK 721
Qy 665 STKVEEDFNLTTEGATKTVTQOQERDIGIIERAPQDQTNMPTSELGRKEDTKSTDSSE 724
Db 722 SV-----PNKALELKNEQTLRADEILP-----SESQKQDYBESSWDS 759
Qy 725 IISVSDTONYECLEPATYQKEIKTTNGKIBES----- 756
Db 760 SLCEVTSQKQVCLPKATHQKEIDKINGKLEESPNDGFLKAPCRMKVSIPTKALELMDMQ 819
Qy 757 -----PEKPSHFEPATQMNSVFNKGLWKNKOTLRA----- 788
Db 820 TFKAEPPEKPSAFPAIEMQSVNKALELQNEQTLRADQMPFSESQKKVENSWSSES 879
Qy 789 -----DSTTLKILDALPSCGERELKKNCEQITA 819
Db 880 LRETVSQKQVCPKATHQKEMDKISGKLEDSTLSKILDTVHSCERARELQKHCEQRTG 939
Qy 820 KMEQMKKFCVLOKELSEAKIKSQLENOKAKWEQELCSVRLPLNOBEERKNVDILKEK 879
Db 940 KMEQMKKFCVLLKKLSAKIKSQLENQKVKEQELCSVRLTLNQBEERKNADILNEK 999
Qy 880 IRP-----EEQRLKLEVKHOLEQTLRIQDIELKSVTSNNOVSHTHSENDLFHENCML 934
Db 1000 IRBELGRIEEQRHRELEVKQLEQALRIQDIELKSVTSNNOVSHTHSENDLFHENCML 1059
Qy 935 KKEIAMLKLEVATLKHOQVKNKYFEDIKILQEKNAELQMTLKLKQKVTVKRASQYREQ 994
Db 1060 KKEIAMLKLEIATLKHOQVKNKYFEDIKILKEKNAELQMTLKLKESLTKRASQYSGQ 1119
Qy 995 LKVLTAENTMLTSKLKE 1011
Db 1120 LKVLIAENTMLTSKLKE 1136
```



```
Publication No. US20030023036A1
GENERAL INFORMATION:
APPLICANT: Houghton, Raymond L.
APPLICANT: Sleath, Paul R.
APPLICANT: Persing, David H.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
FILE REFERENCE: 210121.470C11
CURRENT APPLICATION NUMBER: US/10/076.622
CURRENT FILING DATE: 2002-02-13
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 573
LENGTH: 1349
TYPE: PRT
ORGANISM: Homo sapiens
US-10-076-622-573

Query Match 59.3%; Score 3069; DB 14; Length 1349;
Best Local Similarity 57.3%; Pred. No. 1.8e-152;
Matches 663; Conservative 76; Mismatches 146; Indels 272; Gaps 12;

QY 1 MVATLLSYGAVIEVONKASLTPLLLAIQKRSKQVTEFLLTKNANANAFNSKCTALMLAI 60
Db VVAKLLSHGAVIEVHNKASLTPLLLSITKRSQVIEFLLIKNANANAVNKYKCTALMLAV 165
QY 61 CEGSEIVGMLLOQNVDFAEIDHIGITAERYAAARGVNYTHOOLLEHRLKPKNPQNTNP 120
Db CHGLSEIVGMLLOQNVDFAAADICGVTAHYAVTCGFGHHIHEQIMEYIRKLSKKNHNTNP 225
QY 121 EGTSTGTDPDEAAPLA-----ERTPDATASLLEKTPDAAALIVE 158
Db EGTSTGTDPDEAAPLAERTPDATASLVEKTPDEAAPLVETPDATASLVEKTPDEAAALIVE 285
QY 159 GTSKIOCLKATSGKFPQSTTEPRKILRPTKTSKFSWP----- 200
Db GTSKIOCLKATSGKFPQSAETPREITSKTSKFSWP----- 345
QY 201 -----AKERSKITWEEKTSVKTECVAGTVPNKTEVLEKGTSMIACP 244
Db EIMSPAKTSEKFTWAAGRPRKIAWEKETPVKTCVARTSNKRVLEKGRKMIACP 405
QY 245 TKETSTKATNVDSVSEPIFSLFGRTRIENSQCTKVEEDFNALTKIISAAQNYTCLP 304
Db TKESSTKASAN----- 416
QY 305 DATYOKOIKTINHIEDQMPSESREDEDEYSDWSDGSLFESSAKTQVCIPESMYOKWME 364
Db DQRFPSKQDEDEYSDWSDGSLFESSAKTQVCIPESMYOKWME 460
QY 365 INREVEELPEKPSAKPAVEMQKTPNKAFLKNEQTLRAAQMPSESQKQDDENSWS 424
Db INREVEEPKPSAKPALEMONSPNKAFLKNEQTLRAADPMFPPEKQKDYENSWS 520
QY 425 ESPCTVSKQVYLKPAHQEFDTLSGKLEESPVKDGLLKPTCGRKVSLFNKALELKDOR 484
Db ESLCTVSKQVCLPKAQHQEIDKINGKLEESPNKDGLLKATCGMKVSIPTKALELXOM 580
QY 485 EFTKASPDGGLLKPTCGRKVSLFNKALELKDRETLKASPDNDGLLKPTCGRKVSLPN 544
Db QTFKAEPGKPSAFEP-----ATEMOK-----SVFN 606
QY 545 KALELKORTFFKAAQMFPSKQKDEENSWDFESFLETLNDVCLPKAQHQEFDTLS 604
Db KALELKNEQTLRADELLFSESQKDYENSWSDESLCTVSKQVCLPKAQHQEIDKIN 666
QY 605 GKLEESPDGGLLKPTCGMKISLPNKALELKDRETFKADVSSVSTFSLFGKPTTENSQ 664
Db GKLESPVKDGLLKANGCMKVSIPTKALELMDMQTFKAEPPEK-PSAF-----EPAIEMOK 721
QY 665 STKVEEDFNLTKEGATKVTTCQOERDGIIRAPQDQTNKNKPTSELGKEDTKTSOSE 724
Db PNKALELKNEQTLRADEILP-----SESKQKQYEESSEWSDE 759

RESULT 13
US-10-124-805-573
Sequence 573, Application US/10124805
Publication No. US20030166022A1
GENERAL INFORMATION:
APPLICANT: Houghton, Raymond L.
APPLICANT: Sleath, Paul R.
APPLICANT: Persing, David H.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
FILE REFERENCE: 210121.470C12
CURRENT APPLICATION NUMBER: US/10/124.805
CURRENT FILING DATE: 2002-04-15
NUMBER OF SEQ ID NOS: 627
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 573
LENGTH: 1349
TYPE: PRT
ORGANISM: Homo sapiens
US-10-124-805-573

Query Match 59.3%; Score 3069; DB 14; Length 1349;
Best Local Similarity 57.3%; Pred. No. 1.8e-152;
Matches 663; Conservative 76; Mismatches 146; Indels 272; Gaps 12;

QY 1 MVATLLSYGAVIEVONKASLTPLLLAIQKRSKQVTEFLLTKNANANAFNSKCTALMLAI 60
Db VVAKLLSHGAVIEVHNKASLTPLLLSITKRSQVIEFLLIKNANANAVNKYKCTALMLAV 165
QY 61 CEGSEIVGMLLOQNVDFAEIDHIGITAERYAAARGVNYTHOOLLEHRLKPKNPQNTNP 120
Db CHGLSEIVGMLLOQNVDFAAADICGVTAHYAVTCGFGHHIHEQIMEYIRKLSKKNHNTNP 225
QY 121 EGTSTGTDPDEAAPLA-----ERTPDATASLLEKTPDAAALIVE 158
Db EGTSTGTDPDEAAPLAERTPDATASLVEKTPDEAAPLVETPDATASLVEKTPDEAAALIVE 285
QY 159 GTSKIOCLKATSGKFPQSTTEPRKILRPTKTSKFSWP----- 200
Db GTSKIOCLKATSGKFPQSAETPREITSKTSKFSWP----- 345
QY 201 -----AKERSKITWEEKTSVKTECVAGTVPNKTEVLEKGTSMIACP 244
Db EIMSPAKTSEKFTWAAGRPRKIAWEKETPVKTCVARTSNKRVLEKGRKMIACP 405
QY 245 TKETSTKATNVDSVSEPIFSLFGRTRIENSQCTKVEEDFNALTKIISAAQNYTCLP 304
Db TKESSTKASAN----- 416
QY 305 DATYOKOIKTINHIEDQMPSESREDEDEYSDWSDGSLFESSAKTQVCIPESMYOKWME 364
Db DQRFPSKQDEDEYSDWSDGSLFESSAKTQVCIPESMYOKWME 460
QY 365 INREVEELPEKPSAKPAVEMQKTPNKAFLKNEQTLRAAQMPSESQKQDDENSWS 424
Db INREVEEPKPSAKPALEMONSPNKAFLKNEQTLRAADPMFPPEKQKDYENSWS 520
QY 425 ESPCTVSKQVYLKPAHQEFDTLSGKLEESPVKDGLLKPTCGRKVSLFNKALELKDOR 484
Db ESLCTVSKQVCLPKAQHQEIDKINGKLEESPNKDGLLKATCGMKVSIPTKALELXOM 580
QY 485 EFTKASPDGGLLKPTCGRKVSLFNKALELKDRETLKASPDNDGLLKPTCGRKVSLPN 544
Db QTFKAEPGKPSAFEP-----ATEMOK-----SVFN 606
QY 545 KALELKORTFFKAAQMFPSKQKDEENSWDFESFLETLNDVCLPKAQHQEFDTLS 604
Db KALELKNEQTLRADELLFSESQKDYENSWSDESLCTVSKQVCLPKAQHQEIDKIN 666
QY 605 GKLEESPDGGLLKPTCGMKISLPNKALELKDRETFKADVSSVSTFSLFGKPTTENSQ 664
Db GKLESPVKDGLLKANGCMKVSIPTKALELMDMQTFKAEPPEK-PSAF-----EPAIEMOK 721
QY 665 STKVEEDFNLTKEGATKVTTCQOERDGIIRAPQDQTNKNKPTSELGKEDTKTSOSE 724
Db PNKALELKNEQTLRADEILP-----SESKQKQYEESSEWSDE 759
```



```
Db 346 EIMGPAKETSBEKFTWAAKGRPRKIAWEKKEPTVKTGCVARVTSNKTIVLEKGRSKMIACP 405
Qy 245 TKETSTKASTNDVSSVEPIFSLFCTRTRIENSQCTKVEEDFNLATKIISKAAQNYTCLP 304
Db 406 TKESSTKASAN----- 416
Qy 305 DATYQDKITINHIEDOMPSESKREDEBYSDSGSLFESSAKTQVCIPESMYQKWE 364
Db 417 -----DQFPSESKQDEDEYSCDSRSLFESSAKIQCIPESYQKWE 460
Qy 365 INREVEELPEKPSAFKPAVENQKTPVNPKAFLKNEOTLRAAQMPSESKQKDDRENSWDS 424
Db 461 INREVEEPKPSAFKPAIEMQNSVPNKAFLKNEOTLRAQMPSESKQKDDRENSWDS 520
Qy 425 ESPCETVSQKDVYLPKATHQKEFDLTKGLESVPKGLKPTCGRKVSLPNKALELADR 484
Db 521 ESLCETVSQKDVCLPKATHQKEIDKINGKLESVPNKGSLKATCGMKVSIPTKALELADM 580
Qy 485 ETFKAESPDKGLLKPTCGRKVSLPNKALELADRTELKAESEPDNDGLLKPTCGRKVSLPN 544
Db 581 QTFKAEPKPSAFEP-----ATEMOK-----SVPN 606
Qy 545 KALELKORETFKAAQMPSESKQKDDRENSWDSFLETLQNDVCLPKATHQKEFDLTKS 604
Db 607 KALELKNEOTLRADEILPSESKQKDYENSWDTSLETCETVSQKDVCLPKAAHQKEIDKIN 666
Qy 605 KGLSESPDKGLLKPTCGMKISLPNKALELADRTEFKAEDVSSVESTPSLFGKPTENSQ 664
Db 667 KGLSESPDKGLLKPTCGMKVSIPTKALELADMOTFKAEPPEK-PSAF-----EPAIEMOK 721
Qy 665 STKVEEDFNLTGKATKVTGQOERDIGIIERAPQDQTNKMPSELGRKEDTKSTDSSE 724
Db 722 SV-----PNKALELKNEOTLRADEILP-----SESKQDYESSWDS 759
Qy 725 IISVSDTONYCLPEATYQKEIKTTNGKIEES----- 756
Db 760 SLCETVSQKDVCLPKATHQKEIDKINGKLESSEPDNDGLFKAPCRMKVSIPTKALELMDM 819
Qy 757 -----PEKPSHPEPATEMQNSVPNKGLEWKNKOTLRA----- 788
Db 820 TFKAEPPKPSAFPAIEMQNSVPNKALELKNEOTLRAQMPSESKQKDYENSWDS 879
Qy 789 -----DSTTLISKILDALPSCERGERELKKNCEQITA 819
Db 880 LRETVSQKDVCPKATHQKEMDKISGLEDSTLSKILDTVHSCERARELQKHCEQRTG 939
Qy 820 KMEQKXKFCVLOKELSAKEIKSQLENQKAKWQELCSVPLNQBEEKRNVDILKEK 879
Db 940 KMEQKXKFCVLOKELSAKEIKSQLENQKAKWQELCSVPLNQBEEKRNVDILNEK 999
Qy 880 IRP-----EQLRKKLEVKHOLEOTLRIDQIELSVTSNLNQVSHTHSENDLPHENCM 934
Db 1000 IRELGRIEEOHRELEVKQLEQALRIQDIELKSVESNLNQVSHTHENYLLHENCML 1059
Qy 935 KKEIAMLKLEATLKHQHVKNKYFEDIKIQENKAELOMTLKLQKVTVKRASQYREQ 994
Db 1060 KKEIAMLKLEATLKHQYQENKYNFYEDIKILKEKNAELQMTLKLKESLTKRASQYSGQ 1119
Qy 995 LKVLTAEINTMLTSKLKE 1011
Db 1120 LKVLTAEINTMLTSKLKE 1136
```

RESULT 14

```
US-10-007-805-553
; Sequence 553, Application US/10007805
; Publication No. US20020150581A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
```

```
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margarita
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C10
; CURRENT APPLICATION NUMBER: US/10/007,805
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 593
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 553
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-805-553
```

```
Query Match 45.9%; Score 2374; DB 13; Length 1013;
Best Local Similarity 56.0%; Pred. No. 3.4e-116;
Matches 512; Conservative 61; Mismatches 126; Indels 216; Gaps 9;

Qy 187 LRPTKTSKPSWPAKERSRKITWEEKETSVKTECVAGVTPNKTEVLEKGTSMIACPTK 246
Db 12 MSPAKETSEKFTWAAKGRPRKIAWEKKEPTVKTGCVARVTSNKTIVLEKGRSKMIACP 71
Qy 247 ETSKASTNVDSVSEPIFSLFGTRTIENSQCTKVEEDFNLATKIISKAAQNYTCLPDA 306
Db 72 ESTKASAN----- 80
Qy 307 TYQDKITINHIEDOMPSESKREDEBYSDSGSLFESSAKTQVCIPESMYQKWEIN 366
Db 81 -----DQFPSESKQDEDEYSCDSRSLFESSAKIQCIPESYQKWEIN 126
Qy 367 REVEELPEKPSAFKPAVENQKTPVNPKAFLKNEOTLRAAQMPSESKQKDDRENSWDS 426
Db 127 REVEEPKPSAFKPAIEMQNSVPNKAFLKNEOTLRAQMPSESKQKDYENSWDS 186
Qy 427 PCETVSQKDVYLPKATHQKEFDLTKGLESVPKGLKPTCGRKVSLPNKALELADR 486
Db 187 LCETVSQKDVCLPKATHQKEIDKINGKLESSEPDNDGLLKATCGMKVSIPTKALELADM 246
Qy 487 FKAEPPKPSAFKPAIEMQNSVPNKALELKNEOTLRAQMPSESKQKDYENSWDS 546
Db 247 FKAEPPKPSAFKPAIEMQNSVPNKALELKNEOTLRAQMPSESKQKDYENSWDS 284
Qy 547 LELKDRTEFKAQMPSESKQKDDRENSWDSFLETLQNDVCLPKATHQKEFDLTKSGK 606
Db 285 -----EILPSESKQKDYENSWDSLETCETVSQKDVCLPKAAHQKEIDKINGK 332
Qy 607 LEESPDKGLLKPTCGMKISLPNKALELKNEOTLRAQMPSESKQKDDRENSWDS 666
Db 333 LEGSPDKGLLKANGCMKVSIPTKALELADMOTFKAEPPEK-PSAF-----EPAIEMQNSV 387
Qy 667 KVEEDFNLTGKATKVTGQOERDIGIIERAPQDQTNKMPSELGRKEDTKSTDSBII 726
Db 388 -----PNKALELKNEOTLRADEILP-----SESKQDYESSWDS 425
Qy 727 SVSDTONYCLPEATYQKEIKTTNGKIEES----- 756
Db 426 CETVSQKDVCLPKAAHQKEIDKINGKLESSEPDNDGLFKAPCRMKVSIPTKALELMDMOT 485
Qy 757 -----PEKPSHPEPATEMQNSVPNKGLEWKNKOTLRA----- 788
Db 486 KAEPPKPSAFKPAIEMQNSVPNKALELKNEOTLRAQMPSESKQKDYENSWDS 545
Qy 789 -----DSTTLISKILDALPSCERGERELKKNCEQITAKM 821
Db 546 ETVSQKDVCPKATHQKEMDKISGLEDSTLSKILDTVHSCERARELQKHCEQRTGM 605
Qy 822 EQMKNKFCVLOKELSAKEIKSQLENQKAKWQELCSVPLNQBEEKRNVDILKEKIR 881
```

Db	606	EQMKKFCVLLKKLSEAKEIKSQLENOKVQWQELCSVRLTLNQEBEKRNDILNEKIR	665
Qy	882	P-----EQLRKKLEVKHQLEQTLRIQDIELKSVTSNLSNVSHTHSENDLFHENCMLKK	936
Db	666	EELGRIEEQRKELEVKQLEQALRIQDIELKSVESNLSNVSHTHENENYLLHENCMLKK	725
Qy	937	EIAMLKLEVATLKHQHVKENKYPEDIKILOEKNAELQMTLKLKQKTVTKRASQYREOLK	996
Db	726	EIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKESLTKRASQYSGQLK	785
Qy	997	VLTAENTMLTSKLKE 1011	
Db	786	VLIAENTMLTSKLKE 800	
RESULT 15			
US-10-076-622-553			
; Sequence 553, Application US/10076622			
; Publication No. US20030023036A1			
; GENERAL INFORMATION:			
; APPLICANT: Houghton, Raymond L.			
; APPLICANT: Sleath, Paul R.			
; APPLICANT: Persing, David H.			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY			
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER			
; FILE REFERENCE: 210121.470C11			
; CURRENT APPLICATION NUMBER: US/10/076,622			
; CURRENT FILING DATE: 2002-02-13			
; NUMBER OF SEQ ID NOS: 627			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 553			
; LENGTH: 1013			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-076-622-553			
Query Match 45.9%; Score 2374; DB 14; Length 1013;			
Best Local Similarity 56.0%; Pred No. 3 4e-116;			
Matches 512; Conservative 61; Mismatches 126; Indels 216; Gaps 9;			
Qy	187	LRPRTKTSKFSWPAKERSKITWEEKTSVKTECVAGVTPNKTVELEKGTSGNMIACPTK	246
Db	12	MSPAKETSEKFTWAAKGRPRKIAWEKKETPVKTGCVARVTSNKTVKLEKRSKMIACPTK	71
Qy	247	ETSTKASTNVDSVSEPIESLFGTRTIENSQCTKVEEDFNLATKIISKSAQNYTCLPDA	306
Db	72	ESSTKASAN-----	80
Qy	307	TYQKOIKTINHKIEDQMPFSESKREDEEYSWDSGLPSSAKTQVCIPESMYQKWEIN	366
Db	81	-----DQRFSESKQEBDEEYSCDSRSLPSSAKIQVCIPESYQKWEIN	126
Qy	367	REVEELPEKPSAPKPAVENQKTVPNKAFELKNEQTLRAAQMPFSSSKQKDDSENSWDS	426
Db	127	REVEEPKPSAPKPAIEMQNSVPNKAFELKNEQTLRADPNPFPFSSKQKDYENSWS	186
Qy	427	PCETVSKQVYLPKATHQKFEFTLSGLKLESVPVKQGLLKPTCGRKVSLPNKALELKDR	486
Db	187	LCETVSKQVCLPKATHQKFEIDKINGKLESVPNKQGLLKATCGMKVSIPTKALELKD	246
Qy	487	FKAESPDKGLLKPTCGRKVSLPNKALELKDRTELKAESPDNDGLLKPTCGRKVSLPNKA	546
Db	247	FKAEPKPSAPFAPATQMOKSVPNKALELKNEQTLRAD-----	284
Qy	547	LELKDRTEFFKAAQMPFSSSKQKDDSENSWDFLETLLONDVCLPKATHQKFEFTLSGK	606
Db	285	-----EILPSESQKQDYENSWDTESLCETVSKQKDYENSWDTEFLPAAHQKIDKING	332
Qy	607	LEBSPDKGLLKPTCGMKKISLPNKALELKDRTEFFKAEADVSSVSTFSLFGKPTTENSQ	666
Db	333	LEBSPDKGLLKANCCKMKVSIPTKALELMDQMTFKAEPPEK-PSAF-----EPAIEMQKSV	387
Qy	667	KVEEDFNLTTKBGAATKTVTQOQRDGIIEERAFQDQTNKMPTELGRKEDTKTSDSEII	726

Search completed: February 19, 2005, 02:33:53
Job time : 101.958 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	2372.5	45.9	1095	4	US-09-620-405B-493	Sequence 493, App
2	2372.5	45.9	1095	4	US-09-834-759-493	Sequence 493, App
3	2367	45.8	1002	4	US-09-620-405B-475	Sequence 475, App
4	2367	45.8	1002	4	US-09-604-287A-475	Sequence 475, App
5	2367	45.8	1002	4	US-09-834-759-475	Sequence 475, App
6	2367	45.8	1002	4	US-09-590-751A-475	Sequence 475, App
7	2367	45.8	1002	4	US-09-551-621-475	Sequence 475, App
8	1756.5	34.0	743	4	US-09-620-405B-494	Sequence 494, App
9	1756.5	34.0	743	4	US-09-834-759-494	Sequence 494, App
10	1751	33.8	650	4	US-09-620-405B-469	Sequence 469, App
11	1751	33.8	650	4	US-09-433-826B-469	Sequence 469, App
12	1751	33.8	650	4	US-09-604-287A-469	Sequence 469, App
13	1751	33.8	650	4	US-09-834-759-469	Sequence 469, App
14	1751	33.8	650	4	US-09-590-751A-469	Sequence 469, App
15	1751	33.8	650	4	US-09-551-621-469	Sequence 469, App
16	1542	29.8	512	4	US-09-451-739H-16	Sequence 16, App1
17	1414.5	27.3	445	4	US-09-620-405B-473	Sequence 473, App
18	1414.5	27.3	445	4	US-09-433-826B-473	Sequence 473, App
19	1414.5	27.3	445	4	US-09-604-287A-473	Sequence 473, App
20	1414.5	27.3	445	4	US-09-834-759-473	Sequence 473, App
21	1414.5	27.3	445	4	US-09-590-751A-473	Sequence 473, App
22	1414.5	27.3	445	4	US-09-551-621-473	Sequence 473, App
23	1384	26.8	466	4	US-09-620-405B-472	Sequence 472, App
24	1384	26.8	466	4	US-09-433-826B-472	Sequence 472, App
25	1384	26.8	466	4	US-09-604-287A-472	Sequence 472, App
26	1384	26.8	466	4	US-09-834-759-472	Sequence 472, App
27	1384	26.8	466	4	US-09-590-751A-472	Sequence 472, App

Tue Feb 22 09:49:15 2005

```

170 SKQEDBEYSWDSGLFESSAKIQVCIPEISYQKVMINREVEEPPKPSAFPAIEMQN 229
171 TVPNKAPFELKNEQTLRAAQMPPSSKQKDDSEENSWDSSEPCETVSQKQVYLPKATHQKEP 447
172 SVPNKAPFELKNEQTLRADPMFPESKQKDYENSWDSSELCETVSQKQVCLPKATHQKEI 289
173 DTLGKLEESPVKQGLLKPTCGRKVSIPNKALELDKRETTKKAESPDKGGLLKPTCGRKVS 507
174 DKINGKLEESPNKQGLLKATCMKVSIPNKALELDKRETTKKAESPDKGGLLKPTCGRKVS 349
175 LPNKALELDKRETTKKAESPDKGGLLKATCMKVSIPNKALELDKRETTKKAESPDKGGLLKPTCGRKVS 567
176 VNKALELKNQTLRADPMFPESKQKDYENSWDSSELCETVSQKQVCLPKATHQKEI 375
177 KDDSEENSWDSGLFESSAKIQVCIPEISYQKVMINREVEEPPKPSAFPAIEMQN 627
178 KDYENSWDSSELCETVSQKQVCLPKATHQKEI 435
179 PNKALELDKRETTKKAESPDKGGLLKATCMKVSIPNKALELDKRETTKKAESPDKGGLLKPTCGRKVS 687
180 PNKALELDKRETTKKAESPDKGGLLKATCMKVSIPNKALELDKRETTKKAESPDKGGLLKPTCGRKVS 477
181 QERDIGIERAPQDQTNMPTSELGRKEDTKSTDSSEIISVSDIQNYECLPEATYQKEIK 747
182 NEQTLRADDEILP-----SEKQDYESSWDSSELCETVSQKQVCLPKATHQKEI 528
183 TTNGKIEES-----EPAIEMQKSV-----PNKALELK 773
184 KINGKLEESPDNDGFLKAPCMKVSIPNKALELDKRETTKKAESPDKGGLLKPTCGRKVS 588
185 PNKALELDKRETTKKAESPDKGGLLKATCMKVSIPNKALELDKRETTKKAESPDKGGLLKPTCGRKVS 788
186 PNKALELDKRETTKKAESPDKGGLLKATCMKVSIPNKALELDKRETTKKAESPDKGGLLKPTCGRKVS 648
187 DTTLSKILDALPSCERGLKDNCEQITAKMEQMKNFVCLQKELSEAKIEK 842
188 ISGKLEDSLSKILDTVHSCERARELOKHCEQRTGKMEQMKNFVCLQKELSEAKIEK 708
189 SOLENOKAKWEOLCSVRLPLNQBEEKRNVDILKKEIRP-----EOLRKLKLEVKHOLE 897
190 SOLENOKAKWEOLCSVRLPLNQBEEKRNVDILKKEIRP-----EOLRKLKLEVKHOLE 897
191 SOLENOKAKWEOLCSVRLPLNQBEEKRNVDILKKEIRP-----EOLRKLKLEVKHOLE 897
192 KYFEDIKILQEKNAELQMTLKQKVTTRKASQYREQLKVLTAENTMLTSLKLE 1011
193 KYFEDIKILQEKNAELQMTLKQKVTTRKASQYREQLKVLTAENTMLTSLKLE 882

```

```

RESULT 2
US-09-834-759-493
; Sequence 493, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 493
; LENGTH: 1095
; TYPE: PRT

```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(1095)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-834-759-493

Query Match      45.9%; Score 2372.5; DB 4; Length 1095;
Best Local Similarity 52.2%; Pred. No. 1.9e-156;
Matches 529; Conservative 73; Mismatches 151; Indels 261; Gaps 13;

104 LLEHRLKLPKNQNTNPEGTGTGTPDEAALPABRTPDTAESLLEKTPDEAALVGTSAK 163
24 LLENVISTKINQVSKTE-----YKELLQBFIDNA-----TTNA 58
164 IQCLGKATSGKFEQSTETPRKI-----LRPTKETSFKFWPAKERSRK 207
59 IDELKEC-----FNQIDETLSNVFPMQLIYDSSLCDLFPMSPAKETSEKFTWAAGRPRK 114
208 ITWEEKETSVKTECVAGVTPNKTVELEKTSNNIACPTKETSTKASTNVDSVVEPIFSL 267
115 LAWEKKEKTPVKTGCVARVTSNKTVLEKGRSMIACPTKESSTKASAN-----162
268 FGTRTTSNQCTKVEEDFNLAFTKISKAAQNYTCLPDATYQKDIKTNHKTEDQMPFSE 327
163 -----DQFPFSE 169
328 SKREEDBEYSWDSGLFESSAKIQVCIPEISYQKVMINREVEEPPKPSAFPAIEMQN 387
170 SKQEDBEYSWDSGLFESSAKIQVCIPEISYQKVMINREVEEPPKPSAFPAIEMQN 229
388 TVPNKAPFELKNEQTLRAAQMPPSSKQKDDSEENSWDSSEPCETVSQKQVYLPKATHQKEF 447
230 SVPNKAPFELKNEQTLRADPMFPESKQKDYENSWDSSELCETVSQKQVCLPKATHQKEI 289
448 DTLGKLEESPVKQGLLKPTCGRKVSIPNKALELDKRETTKKAESPDKGGLLKPTCGRKVS 507
290 DKINGKLEESPNKQGLLKATCMKVSIPNKALELDKRETTKKAESPDKGGLLKPTCGRKVS 349
508 LPNKALELDKRETTKKAESPDKGGLLKPTCGRKVSIPNKALELDKRETTKKAESPDKGGLLKPTCGRKVS 567
350 VNKALELKNQTLRADPMFPESKQKDYENSWDSSELCETVSQKQVCLPKATHQKEI 375
568 KDDSEENSWDSGLFESSAKIQVCIPEISYQKVMINREVEEPPKPSAFPAIEMQN 627
376 KDYENSWDSSELCETVSQKQVCLPKATHQKEI 435
628 PNKALELDKRETTKKAESPDKGGLLKPTCGRKVSIPNKALELDKRETTKKAESPDKGGLLKPTCGRKVS 687
436 PNKALELDKRETTKKAESPDKGGLLKPTCGRKVSIPNKALELDKRETTKKAESPDKGGLLKPTCGRKVS 477
688 QERDIGIERAPQDQTNMPTSELGRKEDTKSTDSSEIISVSDIQNYECLPEATYQKEIK 747
478 NEQTLRADDEILP-----SEKQDYESSWDSSELCETVSQKQVCLPKATHQKEI 528
748 TTNGKIEES-----PEKPSHPEPATEMNSV 773
529 KINGKLEESPDNDGFLKAPCMKVSIPNKALELDKRETTKKAESPDKGGLLKPTCGRKVS 588
774 PNKGLWKNKQTLRA-----888
589 PNKALELDKRETTKKAESPDKGGLLKATCMKVSIPNKALELDKRETTKKAESPDKGGLLKPTCGRKVS 648
789 -----DSTTLSKILDALPSCERGLKDNCEQITAKMEQMKNFVCLQKELSEAKIEK 842
649 ISGKLEDSLSKILDTVHSCERARELOKHCEQRTGKMEQMKNFVCLQKELSEAKIEK 708
843 SOLENOKAKWEOLCSVRLPLNQBEEKRNVDILKKEIRP-----EOLRKLKLEVKHOLE 897
709 SOLENOKAKWEOLCSVRLPLNQBEEKRNVDILKKEIRP-----EOLRKLKLEVKHOLE 897
898 QTLRIQDIELKSVTSLNQNVSHTHSENDLFHENCMLKKEIAMLKLEIATLKHQVQKEN 957

```

Query Match	45.8%;	Score 2367;	DB 4;	Length 1002;
Best Local Similarity	55.8%;	Pred. NO. 4.1e-156;		
Matches	511;	Conservative 62;	Mismatches 126;	Indels 216;
Gaps	9			
Qy	187	LRPTEKTESEKSWPAKERSRKITWEEKETSVKTECVAGVTPNKTVELEKGTSMNIACPTK	246	
Db	1	MSPAKETSEKFTWAAKGPRKIAWKEKSTPVKTCVAVRTSNKTKVLEKGRSKMIACPTK	60	
Qy	247	ETSTKASTNVDSVVEPIFSLFGTTRTENSQCTKVEEDFNLATKIISKSAQNYTCLPDA	306	
Db	61	ESSTKASAN	69	
Qy	307	TYQDKDITINHKI EQMPPSSKREDEEYSWDSGLSFESSAKTOVCIPESMYQKWEIN	366	
Db	70	-----DQRPSSKQBEDEEYSCDSRSLFESSAKIQVCIPESIQKWEIN	115	
Qy	367	REVELPEKPSAFKPAVEMQKTVPNKAPFELKNEOTLRAAQMPSSKQKDDSENSWDS	426	
Db	116	REVEEPPKPSAFKPAIEMQNSVPNKAPFELKNEQTLRADPMPSPSSKQDYENSWS	175	
Qy	427	PCETVSQKDVVLPAKTHQKEFDTLISGLSESPVKDGLLKPTCGRKVSLPNKALELKDRET	486	
Db	176	LCETVSQKDVCLPAKTHQKEIDKINGKLEESPNKDGLLKATCGMKVSIPTKALELKDMQT	235	
Qy	487	FKAESPDKDGLLKPTCGRKVSLPNKALELKORETLKAESPNDGLLKPTCGRKVSLPNKA	546	
Db	236	FKAEPGPKPSAFEPATENMQKSVPNKALEKNEQTLRAD	273	
Qy	547	LELKDRETFFKAAQMPSPSSKQKDDSENSWDPESFLETLIQNDVCLPAKTHQKEFDTLGSK	606	
Db	274	-----EILPSSKQKDYESSWDSSELCETVSQKDVCLPAKHQKEIDKINGK	321	
Qy	607	LEESPDKDGLLKPTCGMKISLPNKALELKDRETFKAEADVSSVESTFSFLGKPTTENSQST	666	
Db	322	LEGSPVKDGLLKANCQMKVSIPTKALELMDMQTFKAEPEK-PSAF----	376	
Qy	667	KVEEDFNLTTEKATKTVTGOQERDIGIIERAPQDQTNKQPTSELGRKEDTKSTSDSEII	726	

```
Db 70 -----DQFPSESKQDEDEYSCDSRSLPSESSAKIQVCIPESYQKVMEN 115
Qy 367 REVELPEKPSAFKPAVMQKTVPNKAFELKNEQTLRAAQMPFSESKQKDEENSWSES 426
Db 116 REVEPPKPSAFKPAIEMQNSVNPKNKAFELKNEQTLRADPMFPPESESKQKDEENSWSES 175
Qy 427 PCETVSQKDVLPKATHQKEEDTLGKLEESPVKDGKLLKPTCGRKVSLPNKALELKDR 486
Db 176 LCETVSQKDVLPKATHQKEIDKINGKLEESPNKDGKLLKATCGMKVSIPTKALELKMOT 235
Qy 487 FKAESPDKDGLLKPTCGRKVSLPNKALELKDRITLKAESPNDGKLLKPTCGRKVSLPNKA 546
Db 236 FKAEPKPSAFKPAIEMQNSVNPKNKAFELKNEQTLRAD----- 273
Qy 547 LELKDRITFKAQMPFSESKQKDEENSWSESFLETLQNDVCLPKATHQKEEDTLGK 606
Db 274 -----EILPSESKQKDYBESSWSESLETCETVSQKDVLPKAXHQKEIDKINGK 321
Qy 607 LEESPKDGLLKPTCGMKISLPNKALELKDRITFKAEDVSSVESTFSLFGKPTTENSOST 666
Db 322 LEGSPVKDGLLKANGCMKVSIPKALELMDQMTFKAEPPEK-PSAF-----EPAIEMQKSV 376
Qy 667 KVEDFNLTKEGATKVTGQOERDIIIRAPODQTNKMTSELGRKEDTKSDSII 726
Db 377 -----PNKALELNEQTLRADEILP-----SESKQKDYBESSWSES 414
Qy 727 SVSDTONYCELPKATHQKEIDKINGKLEESPNKALELNEQTLRAD----- 756
Db 415 CETVSQKDVLPKAXHQKEIDKINGKLEESPNKALELNEQTLRAD----- 474
Qy 757 -----PEKPSHFEPAIEMQNSVNPKNKALELNEQTLRAD----- 788
Db 475 KAEPPEKPSAFKPAIEMQNSVNPKNKALELNEQTLRADQMPFSESKQKDEENSWSES 534
Qy 789 -----DSTTLKILDALPSCERGLKKNQCEITAKM 821
Db 535 ETVSQKDVCPKATHQKEIDKINGKLEESPNKALELNEQTLRADQMPFSESKQKDEENSWSES 594
Qy 822 EQMKHFCVLQKELSEAKEIKSQLENQKAKWEQELCSVRLPLNQVEEKRRNVNLIKIR 881
Db 595 EQMKHFCVLQKELSEAKEIKSQLENQKAKWEQELCSVRLPLNQVEEKRRNVNLIKIR 654
Qy 882 P-----EEQLRKLEVKHOLEQTLRIQDIELKSVTSNINQVSHTHSENDFHENCMLKK 936
Db 937 EIAMLKLEVATLKHQVQKENVKYPEDIKILOKNAELQMTLKLKOKVTVKASQYREOLK 996
Db 715 EIAMLKLEIATLKHQVQKENVKYPEDIKILOKNAELQMTLKLKBESELTVKASQYSGOLK 774
Qy 997 VITAENTMLTSKLKE 1011
Db 775 VLIAENTMLTSKLKE 789
```

RESULT 5

```
US-09-834-759-475
; Sequence 475, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 21021.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
; US-09-834-759-475
```

```
Query Match 45.8%; Score 2367; DB 4; Length 1002;
Best Local Similarity 55.8%; Pred. No. 4.1e-156;
Matches 511; Conservative 62; Mismatches 126; Indels 216; Gaps 9;

Qy 187 LRPTKETSEKPSAPKERSRKITWEKETSVKTECVAGVTPNKTEVLEKGTNNIACTPK 246
Db 1 MSPAKETSEKFTWAAKGRPRKIAWEKETPVKTCVARVTSNKTVKLEGRSRMIACTPK 60
Qy 247 ETSTKASTNVDSVSEPIFSLFGTRTTIENSQCTKVEEDFNLATKIISKAAQNYTCLPDA 306
Db 61 ESSTKASAN----- 69
Qy 307 TYQKDIKTINHIEDQMPFSESKREDEBYSWDSGSLFESSAKTQVCIPESMYQKVMEN 366
Db 70 -----DQFPSESKQDEDEYSCDSRSLFESSAKTQVCIPESYQKVMEN 115
Qy 367 REVELPEKPSAFKPAVMQKTVPNKAFELKNEQTLRAAQMPFSESKQKDEENSWSES 426
Db 116 REVEPPKPSAFKPAIEMQNSVNPKNKAFELKNEQTLRADPMFPPESESKQKDEENSWSES 175
Qy 427 PCETVSQKDVLPKATHQKEEDTLGKLEESPVKDGKLLKPTCGRKVSLPNKALELKDR 486
Db 176 LCETVSQKDVLPKATHQKEIDKINGKLEESPNKDGKLLKATCGMKVSIPTKALELKMOT 235
Qy 487 FKAESPDKDGLLKPTCGRKVSLPNKALELKDRITLKAESPNDGKLLKPTCGRKVSLPNKA 546
Db 236 FKAEPKPSAFKPAIEMQNSVNPKNKALELNEQTLRAD----- 273
Qy 547 LELKDRITFKAQMPFSESKQKDEENSWSESFLETLQNDVCLPKATHQKEEDTLGSK 606
Db 274 -----EILPSESKQKDYBESSWSESLETCETVSQKDVLPKAXHQKEIDKINGK 321
Qy 607 LEESPKDGLLKPTCGMKISLPNKALELKDRITFKAEDVSSVESTFSLFGKPTTENSOST 666
Db 322 LEGSPVKDGLLKANGCMKVSIPKALELMDQMTFKAEPPEK-PSAF-----EPAIEMQKSV 376
Qy 667 KVEDFNLTKEGATKVTGQOERDIIIRAPODQTNKMTSELGRKEDTKSDSII 726
Db 377 -----PNKALELNEQTLRADEILP-----SESKQKDYBESSWSES 414
Qy 727 SVSDTONYCELPKATHQKEIDKINGKLEESPNKALELNEQTLRAD----- 756
Db 415 CETVSQKDVLPKAXHQKEIDKINGKLEESPNKALELNEQTLRADQMPFSESKQKDEENSWSES 474
Qy 757 -----PEKPSHFEPAIEMQNSVNPKNKALELNEQTLRAD----- 788
Db 475 KAEPPEKPSAFKPAIEMQNSVNPKNKALELNEQTLRADQMPFSESKQKDEENSWSES 534
Qy 789 -----DSTTLKILDALPSCERGLKKNQCEITAKM 821
Db 535 ETVSQKDVCPKATHQKEIDKINGKLEESPNKALELNEQTLRADQMPFSESKQKDEENSWSES 594
Qy 822 EQMKHFCVLQKELSEAKEIKSQLENQKAKWEQELCSVRLPLNQVEEKRRNVNLIKIR 881
Db 595 EQMKHFCVLQKELSEAKEIKSQLENQKAKWEQELCSVRLPLNQVEEKRRNVNLIKIR 654
Qy 882 P-----EEQLRKLEVKHOLEQTLRIQDIELKSVTSNINQVSHTHSENDFHENCMLKK 936
Db 937 EIAMLKLEVATLKHQVQKENVKYPEDIKILOKNAELQMTLKLKOKVTVKASQYREOLK 996
Db 715 EIAMLKLEIATLKHQVQKENVKYPEDIKILOKNAELQMTLKLKBESELTVKASQYSGOLK 774
Qy 997 VITAENTMLTSKLKE 1011
Db 775 VLIAENTMLTSKLKE 789
```

Db 715 EIALMLKLEIATLKHQYQKQENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLK 774
Qy 997 VLTAENTMLTSKLKE 1011
Db 775 VLIAENTMLTSKLKE 789

RESULT 6

US-09-590-751A-475
; Sequence 475, Application US/09590751A
; Patent No. 6756477
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C6
; CURRENT APPLICATION NUMBER: US/09/590,751A
; CURRENT FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-590-751A-475

Query Match 45.8%; Score 2367; DB 4; Length 1002;
Best Local Similarity 55.8%; Pred. No. 4.1e-156;
Matches 511; Conservative 62; Mismatches 126; Indels 216; Gaps 9;
Qy 187 LRPKTSEKFSWPAKERSRKITWEEKETSVKTECVAGVTPNKTEVLEKGTNNMIACPTK 246
Db 1 MSPAKETSEKFTWAAKGRPKIAWEKETPVKTGCVARVTSNKTVKLEKGRSKMIACPTK 60
Qy 247 ETSTKASTNVDSVVEPIFSLFGTRTIENSQCTKVBEEDFNLATKIISKSAQNYTCLPDA 306
Db 61 ESSTKASAN----- 69

Qy 307 TYQKDIKTINHKEIDQMPPESEKDEEYSDWSGSLFESSAKTQVCIPESMYOKWMEIN 366
Db 70 -----DQRPPESEKQDEEYSDWSGSLFESSAKTQVCIPESMYOKWMEIN 115
Qy 367 REVELPEKPSAFKPAVMQKTPVNPKAPELKNQOTLRAAOMFPSESKQKODEENSWSDES 426
Db 116 REVEPEPKPSAFKPAEMQNSVNPKAPELKNQOTLRAADNFPPESEKQKODEENSWSDES 175

Qy 427 PCETVSKQDVLPKATHQKEFTDLSGKLEESPVKDGGLKPTCGRKVSLPNKALELKDRET 486
Db 176 LCETVSKQDVLPKATHQKEIDKINGKLEESPNKDGGLKATCGMKVSIPTKALELKDQMT 235

Qy 487 FKASPDKGLLTKTCGRKVSPLNKALELKDRETILKAESPNDGLLTKPTCGRKVSLPNKA 546
Db 236 FKAEPPGKPSAFEPATEMOKSVNPNAELKNEQTLRAD----- 273

Qy 547 LELKDRETFKAAOMFPSESKQKODEENSWSDESLETLLQNDVCLPKATHQKEFTDLSGK 606
Db 274 -----ELPSESKQKODEENSWSDES LCETVSKQDVLPKAXHQKEIDKINGK 321

Qy 607 LEESPDKGLLTKTCGRKVSPLNKALELKDRETFFKAEDVSSVESTFSLFGKPTTENSQST 666
Db 322 LEGSPVKDGLLKCANGMKVSIPTKALELMDQMTFKAEPPK-PSAF-----EPATEMOKSV 376

Qy 667 KVEEDFNLTTKGATKTVTGOQERDIIIRAPQDQNNKMTSELGKREKDTKSDSEII 726
Db 377 -----PNKALELQNTLRADELTP-----SESKQKDYESSWSDES 414

Qy 727 SVSTQNYECLPEATYQKEIKTTNGKIEES----- 756
Db 415 CETVSKQDVCLPKAXHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQMTF 474
Qy 757 -----PEKPSHEPEATEMOKSVNPNAELKNEQTLRADQMPPESEKQKODEENSWSDES 788
Db 475 KAEPPEKPSAFEPATEMOKSVNPNAELKNEQTLRADQMPPESEKQKODEENSWSDES 534
Qy 789 -----DSTTSLKILDALPSCERGLKNDCEQITAKM 821
Db 535 ETVSKQDVCPKATHQKEMDKISGLESTSLKILDTVHSCERARELQKHCEORTGKM 594
Qy 822 EQMKNKFCVLQKELSEAKEIKSOLQENKAKWEQELCSVRLFLNQBEERKRRNVLDLKEKIR 881
Db 595 EQMKKFCVLKELSEAKEIKSOLQENKAKWEQELCSVRLFLNQBEERKRRNADILNEKIR 654
Qy 882 P-----BEQARKKLEVKHOLEQTLRIQDIELKSVTSNLSNQVSHTHESNDLFHENCMLKK 936
Db 655 EELGRIBEOHRRKELEVKQLEQALRIQDIELKSVTSNLSNQVSHTHESNDLFHENCMLKK 714
Qy 937 BIAMKLEVATLKHQVQKQENKYPEDIKILKEKNAELQMTLKLKQKTVTKRASQYSGQLK 996
Db 715 BIAMKLEIATLKHQYQKQENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLK 774
Qy 997 VLTAENTMLTSKLKE 1011
Db 775 VLIAENTMLTSKLKE 789

RESULT 7

US-09-551-621-475
; Sequence 475, Application US/09551621
; Patent No. 6825175
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C5
; CURRENT APPLICATION NUMBER: US/09/551,621
; CURRENT FILING DATE: 2000-04-17
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-551-621-475

Query Match 45.8%; Score 2367; DB 4; Length 1002;
Best Local Similarity 55.8%; Pred. No. 4.1e-156;
Matches 511; Conservative 62; Mismatches 126; Indels 216; Gaps 9;

Qy 187 LRPKTSEKFSWPAKERSRKITWEEKETSVKTECVAGVTPNKTEVLEKGTNNMIACPTK 246
Db 1 MSPAKETSEKFTWAAKGRPKIAWEKETPVKTGCVARVTSNKTVKLEKGRSKMIACPTK 60

Qy 247 ETSTKASTNVDSVVEPIFSLFGTRTIENSQCTKVBEEDFNLATKIISKSAQNYTCLPDA 306
Db 61 ESSTKASAN----- 69

Qy 307 TYQKDIKTINHKEIDQMPPESEKDEEYSDWSGSLFESSAKTQVCIPESMYOKWMEIN 366
Db 70 -----DQRPPESEKQDEEYSDWSGSLFESSAKTQVCIPESMYOKWMEIN 115


```
QY 367 REVELPEKPSAFKPAVEMQKTPVNFPAFELKNEQTLRAAQMPFSESQKQDEENSWSES 426
Db 116 REVEPPPKPSAFKPALEIMNSVFNPAFELKNEQTLRADPMFPFSESQKQDEENSWSES 175
QY 427 PCEIVSQKDVLPKATHQKEFDTLGKLEESPVKDGLLKPTCGRKVSLPNKALELKDBET 486
Db 176 LCETVSQKDVCLPKATHQKEIDKINGKLEESPNKDGLLKATCGMKVSIPTKALELDMQT 235
QY 487 FKAESPDKDGLLKPTCGRKVSLPNKALELKDBETLKAESPNDGLLKPTCGRKVSLPNKA 546
Db 236 FKAEPFGPSAFEPATEMOKSVFNKALELKNEQTLRAD----- 273
QY 547 LELKDRFTKAAQMPFSESQKQDEENSWDFESFLETLQNDVCLPKATHQKEFDTLGK 606
Db 274 -----EILPSSSKQDYESSWDSSELCTVSQKDVCLPKAXHQKEIDKINGK 321
QY 607 LEESPDKGLLKPTCGMKISLPNKALELKDBETFKAEVSVSVESFSLFGKPTTENSOST 666
Db 322 LEGSPVKGGLKANGCMKVSIPTKALELMDMQTFKAEPPEK-PSAF-----EPAIEMOKSV 376
QY 667 KVEEDFNLTKEGATKVTGTQOERDGIIRAPQDQTNKMPSTSELGRKEDTKSTSDSII 726
Db 377 -----PNKALELKNEQTLRADEILP-----SESKQDYESSWDSSEL 414
QY 727 SVSDTQNYECLPEATYQKEIKTTNGKIBES----- 756
Db 415 CETVSQKDVCLPKAXHQKEIDKINGKLEESPNNDGFLKAPCRMKVSIPTKALELMDMQTF 474
QY 757 -----PEKSHPEPATEMQNSVFNKLEWKNQTLRA----- 788
Db 475 KAEPPEKPSAFEPATEMOKSVFNKALELKNEQTLRADQMPFSESQKQKVEENSWDSSEL 534
QY 789 -----DSTTLISKILDALPSCERGERELKDNCEQITAKM 821
Db 535 ETVSQKDVCPKATHQKEMDKISGLESTSLKILDTVHSCERARELQKHCEQRTGM 594
QY 822 EQMKNKFCVLQKELSEAKIQLBNQAKMEQELCSVRLPLNQBEEKRRNVNLIKIR 881
Db 595 EQMKKFCVLKKLSEAKIQLBNQAKMEQELCSVRLPLNQBEEKRRNADILNEKIR 654
QY 882 P-----EQRLKLEVLKQLEQTLRIQDIELKSVTSLNQNQVSHTHESNDLPHENCMKK 936
Db 655 BELGRIEQLKLEVLKQLEQTLRIQDIELKSVTSLNQNQVSHTHENENYLLHENCMKK 714
QY 937 ETAMKLEVALTKHQVKNKYFEDIKLOEKNAELQWTLKIKOKTVTKRASQVREQLK 996
Db 715 ETAMKLEVALTKHQVKNKYFEDIKILKEKNAELQWTLKILKEESLTKRASQVSGQLK 774
QY 997 VLTAENTMLTSKLE 1011
Db 775 VLTAENTMLTSKLE 789
```

RESULT 8

US-09-620-405B-494
Sequence 494, Application US/09620405B
Patent No. 6528054

GENERAL INFORMATION:

APPLICANT: Jiang, Yuqiu
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
FILE REFERENCE: 210121.470C8
CURRENT APPLICATION NUMBER: US/09/620,405B
CURRENT FILING DATE: 2000-07-20
NUMBER OF SEQ ID NOS: 495
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 494
LENGTH: 743

TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variant
LOCATION: (1)...(743)
OTHER INFORMATION: Xaa = Any amino acid
US-09-620-405B-494

Query Match 34.0%; Score 1756.5; DB 4; Length 743;
Best Local Similarity 46.0%; Pred. No. 7.8e-114;
Matches 404; Conservative 67; Mismatches 140; Indels 267; Gaps 13;
QY 104 LLEHRLKPNQPNWTEGTSSTGTTPDEAAPLAERTPTDAESLLEKTPDEAARLVEGTSK 163
Db 24 LLENVISKTINPQVSKTE-----YKELLOEFIDDNA-----TTNA 58
QY 164 IQCLGKATSGKFEOSTETPRKI-----LRPTKETSEKTSWPAKERSRK 207
Db 59 IDELKEC-----FLNQTDTLSNVFVFMQLIYDSSLCOLFMSPAKETSEKFTWAAKGRPK 114
QY 208 ITWEKETSVKTECVAGVTPNKTVELEKGTISNMIACTKETSTKASTNVDVSSVEPIFSL 267
Db 115 IANEKKEITPVKTGCVARVTSNKTVELEKGRSKMIACTPKSSSTKASAN----- 162
QY 268 FGTRTIENSQCTKVEEDFNLATKIIISKAAQNYTCLPDATYQKDIKTINHKBIDQMPSE 327
Db 163 -----DQRFPE 169
QY 328 SKREDEYSWDSGLSFESSAKTOVCIPESMYQKVINREVELPEKPSAFKPAVEMOK 387
Db 170 SKQEBEYSCDSRSLFESSAKIQVCIPESYQKVINREVEPPPKPSAFKPAIEMQN 229
QY 388 TVPNKAFELKNEQTLRAQMPFSESQKQDEENSWDSSELCTVSQKDVCLPKATHQKEF 447
Db 230 SVPNKAFELKNEQTLRADPMFPFSESQKQDYENSWDSSELCTVSQKDVCLPKATHQKEI 289
QY 448 DTLGKLEESPVKDGLLKPTCGRKVSLPNKALELKDBETFKAESPDGKLLKPTCGRKVS 507
Db 290 DKINGKLEESPNKDGLLKATCGMKVSIPTKALELMDMQTFKAEPPEKPSAFEPATEMOKS 349
QY 508 LPNKALELKDBETLKAESPNDGLLKPTCGRKVSLPNKALELKDBETFKAAQMPFSESQK 567
Db 350 VFNKALELKNEQTLRAD-----EILPSESQK 375
QY 568 KQDEENSWDFESFLETLQNDVCLPKATHQKEFDTLGKLEESPDGKLLKPTCGMKISL 627
Db 376 KYEESWDSSELCTVSQKDVCLPKAXHQKEIDKINGKLEESPVKDGLLKANGCMKVS 435
QY 628 PNKALELKDBETFKAEVSVSVESFSLFGKPTTENSQSTKVEEDFNLTKEGATKVTGQ 687
Db 436 PTKALELMDMQTFKAEPPEK-PSAF-----EPAIEMOKSV-----PNKALELK 477
QY 688 QERDIGIIRAPQDQTNKMPSTSELGRKEDTKSTSDSIIISVSDTQNYECLPEATYQKEIK 747
Db 478 NEQTLRADEILP-----SESKQDYESSWDSSELCTVSQKDVCLPKAXHQKEID 528
QY 748 TTNGKIBES-----PEKSHPEPATEMQNSV 773
Db 529 KINGKLEESPNNDGFLKAPCRMKVSIPTKALELMDMQTFKAEPPEKPSAFEPATEMOKSV 588
QY 774 PNKLEWKNQKOTLRA----- 788
Db 589 PNKALELKNEQTLRADQMPFSESQKQKVEENSWDSSELRETVSQKDVCPKATHQKEMDK 648
QY 789 -----DSTTLISKILDALPSCERGERELKDNCEQITAKMEQKMKFCVLQKELSEAKIK 842
Db 649 ISGKLESTSLKILDTVHSCERARELQKHCEQRTGMQKMKFCVLKKLSEAKIK 708
QY 843 SOLENNQAKWEQELCSVRLPLNQBEEKRRNVNLIKIR 880
Db 709 SOLENNQAKWEQELCSVRF-----LTLMKMI 735

RESULT 13
US-09-834-759-469
; Sequence 469, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.

Db 475 KAEPPKPSAFPALEMQKSVNPKALELKNQTLRADQMPPSESKQKXVEENSWDSLSR 534
Qy 789 -----DSTLSKILDALPSCERGERELKKNCEQITAKM 821
Db 535 ETVSQKDVCPKATHOKEMDKISGKLEDSTLSKILDVHSCERARELQKDHCEQRTGKM 594
Qy 822 EQMKNKFCVLQKELSEAKEIKSQLENQAKWEOELCSVRLPLNQEBEERNNVDILKEKI 880
Db 595 EQMKKFCVLKXJSEAKEIKSQLENQKVKWEQELCSVRF-----LTLMKMKI 642

RESULT 14
US-09-590-751A-469
; Sequence 469, Application US/09590751A
; Patent No. 6756477
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C6
; CURRENT APPLICATION NUMBER: US/09/590,751A
; CURRENT FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(650)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-590-751A-469

Query Match 33.8%; Score 1751; DB 4; Length 650;
Best Local Similarity 49.6%; Pred. No. 1.6e-113;
Matches 386; Conservative 56; Mismatches 115; Indels 222; Gaps 9;

Qy 187 LRPTKETSKEFSPWPAKERSRKITWEEKTSVKTECVAGVTNPKTEVLEKGTSGNMIACTPK 246
Db 1 MSPAKETSKEFTWAAGRPRIANEKKEPTVKTGCVARVTSNKTVKLEGRSKMIACTPK 60
Qy 247 ETSTKASTNVDSVSSVEPIPSLFGTRTIENSQCTKVEEDFNLAATKIISKSAQAQNYTCLPDA 306
Db 61 ESSTKASAN----- 69
Qy 307 TYQDKITINHIEDOMPSESKREDEEYSWDSGLPSSAKTOVCIPESMYOKWMEIN 366
Db 70 -----DQRPPESEKQEBEYSCDSRLPSSAKIOVCIPESYQKWEIN 115
Qy 367 REVEELPEKPSAFKPAVENQKTPVNAKAFELKNEQTLRAAQMPPESSKQKDEENSWDSSES 426
Db 116 REVEEPKPSAFKPAENQNSVNPNAKAFELKNEQTLRADPMFPPESSKQKDEENSWDSSES 175
Qy 427 PCETVSQKDVYLPKATHQKEFTLQGLKEESPVKDGLLKPTCCGRKVSGLPNKALELKDRET 486
Db 176 LCETVSQKDVCLPKATHQKEIDKINGKLEESPNKDGLLKATCGMKVSIPTKALELMDQTF 235
Qy 487 FKAESPDGGLLKPTCCGRKVSGLPNKALELKDRETLEKAEKSPDNDGLLKPTCCGRKVSGLPNKA 546
Db 236 FKAEPPEKPSAFEPATEMOMKSVNPKALELKNQTLRAD----- 273
Qy 547 LELKDRETFKAAQMPPESSKQKDEENSWDFSEFTLLQNDVCLPKATHQKEFTLQGLK 606
Db 274 -----EILPSESKQKDEENSWDSSELCTETVSQKDVCLPKAXHQKEIDKINGK 321
Qy 607 LEESPDKDGLLKPTCCGMKISLPNAKALELKDRETFKAEADVSSVESFTSLFGKPTTENSQST 666
Db 116 REVEEPKPSAFKPAENQNSVNPNAKAFELKNEQTLRADPMFPPESSKQKDEENSWDSSES 175

Db 322 LEGSPVKDGLLKANGCMKVSIPTKALELMDQTFKAEPPK-PSAF-----EPAIEMQKSV 376
Qy 667 KVEEDFNLTTEGATKTVTGOQERDGIITERAPQDOTNKMPTSELGRKEDTKTSDSEII 726
Db 377 -----PNKALELKNQTLRADEILP-----SEKQKDYESSWDSSEL 414
Qy 727 SVSDTQNYECLPEATYQKEIKTTNGKIEBS----- 756
Db 415 CETVSQKDVCLPKAXHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQTF 474
Qy 757 -----PKPSHFEPATEMOMKSVNPKALELKNQTLRADQMPPSESKQKXVEENSWDSSES 788
Db 475 KAEPPKPSAFPALEMQKSVNPKALELKNQTLRADQMPPSESKQKXVEENSWDSSES 534
Qy 789 -----DSTLSKILDALPSCERGERELKKNCEQITAKM 821
Db 535 ETVSQKDVCPKATHOKEMDKISGKLEDSTLSKILDVHSCERARELQKDHCEQRTGKM 594
Qy 822 EQMKNKFCVLQKELSEAKEIKSQLENQAKWEOELCSVRLPLNQEBEERNNVDILKEKI 880
Db 595 EQMKKFCVLKXJSEAKEIKSQLENQKVKWEQELCSVRF-----LTLMKMKI 642

RESULT 15
US-09-551-621-469
; Sequence 469, Application US/09551621
; Patent No. 6825175
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C5
; CURRENT APPLICATION NUMBER: US/09/551,621
; CURRENT FILING DATE: 2000-04-17
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (310)
; NAME/KEY: Xaa = Any Amino Acid<221> unsure
; LOCATION: (429)
; NAME/KEY: Xaa = Any Amino Acid<221> unsure
; LOCATION: (522)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-551-621-469

Query Match 33.8%; Score 1751; DB 4; Length 650;
Best Local Similarity 49.6%; Pred. No. 1.6e-113;
Matches 386; Conservative 56; Mismatches 115; Indels 222; Gaps 9;

Qy 187 LRPTKETSKEFSPWPAKERSRKITWEEKTSVKTECVAGVTNPKTEVLEKGTSGNMIACTPK 246
Db 1 MSPAKETSKEFTWAAGRPRIANEKKEPTVKTGCVARVTSNKTVKLEGRSKMIACTPK 60
Qy 247 ETSTKASTNVDSVSSVEPIPSLFGTRTIENSQCTKVEEDFNLAATKIISKSAQAQNYTCLPDA 306
Db 61 ESSTKASAN----- 69
Qy 307 TYQDKITINHIEDOMPSESKREDEEYSWDSGLPSSAKTOVCIPESMYOKWMEIN 366
Db 70 -----DQRPPESEKQEBEYSCDSRLPSSAKIOVCIPESYQKWEIN 115
Qy 367 REVEELPEKPSAFKPAVENQKTPVNAKAFELKNEQTLRAAQMPPESSKQKDEENSWDSSES 426
Db 116 REVEEPKPSAFKPAENQNSVNPNAKAFELKNEQTLRADPMFPPESSKQKDEENSWDSSES 175

```
Qy 427 PCETVSQKDVLPKATHQKEPDTLSGKLEESFVKDGLLKPTCGRKVSIPNKALELKORET 486
Db 176 LCETVSQKDVCLPKATHQKEIDKINGKLEESFNKDGLLKATCGMKVSIPTKALELKDQMT 235
Qy 487 FKAESPDKDGLLKPTCGRKVSIPNKALELKORETLKAESEPDNDGLLKPTCGRKVSIPNKA 546
Db 236 FKAESPFGKPSAFEPATEMOKSVNKALELKNEQTLRAD----- 273
Qy 547 LELKDRETFKAAOPFPSSKQKDEENSWDFESFLETLQNDVCLPKATHQKEPDTLSGK 606
Db 274 -----EILPSESKQKDYEESSWDSSELCETVSQKDVCLPKAXHOKHEIDKINGK 321
Qy 607 LEESPDKDLKPTCGMKISIPNKALELKORETFKAEDVSVSESTFSLFGKPTTENSQST 666
Db 322 LEGSPVKDGLLKANGCMKVSIPNKALELMDQTFKAEPPEK-PSAF---EPAIEMQKSV 376
Qy 667 KVEDFNLTTEGATKTVTGOERDIGIIERAPQDOTNKMPTSELGRKEDTKSTDSEII 726
Db 377 -----PNKALELKNEQTLRADEILP-----SESKQKDYEESSWDSSEL 414
Qy 727 SVSDTONYECLPEATYQKEIKTTNGKIBES----- 756
Db 415 CETVSQKDVCLPKAXHOKHEIDKINGKLEESPDNDGFLKAPCRMKVSIPNKALELMDQTF 474
Qy 757 ----PEKPSHFEPATEMOKSVNKALELKNQTLRADQMPFSESKQKXVENSWSLSR 788
Db 475 KAEPPEKPSAFEPALIMQKSVNKALELKNQTLRADQMPFSESKQKXVENSWSLSR 534
Qy 789 -----DSTLSKILDALPSCERGRELKONCEQITAKM 821
Db 535 ETVSQKDVCPKATHQKEMDKISGKLEDSTLSKILDVHSCERARELQKHCEQRTCKM 594
Qy 822 EOMKNKFCVLOKELSEAKEIKSOLENOKAKWEQELCSVRLPLNOEERKRRNVILKEKI 880
Db 595 EOMKKKFCVLKKKUSEAKEIKSOLENOKAKWEQELCSVRF-----LTLMKMKI 642
```

Search completed: February 19, 2005, 02:13:36
Job time : 33.5342 secs

the eye blank (uspto)

Result No.	Score	Query			DB	ID	Description
		Match	Length	%			
1	5173	100.0	1011	4	AAB84703	Aab84703 Amino aci	
2	5173	100.0	1011	7	ADJ69751	Adj69751 Human hea	
3	5173	100.0	1239	6	ABU37789	Abu37789 Human tum	
4	5173	100.0	1239	6	ADL93329	Adl93329 Human bre	
5	5156	99.7	1235	6	ABR47547	AbR47547 Breast ca	
6	3075	59.4	1341	4	AAB84702	Aab84702 Amino aci	
7	3075	59.4	1341	5	ABJ05537	Abj05537 Breast ca	
8	3075	59.4	1341	6	ABU37784	Abu37784 Human tum	
9	3075	59.4	1341	6	ABR47548	AbR47548 Breast ca	
10	3075	59.4	1341	7	ADL93327	Adl93327 Human bre	
11	3069	59.3	1349	6	ABU37788	Abu37788 Human tum	
12	3069	59.3	1349	7	ADL93325	Adl93325 Human bre	
13	2374	45.9	1013	6	ABU37783	Abu37783 Human tum	
14	2374	45.9	1013	7	ADL93315	Adl93315 Human bre	
15	2372.5	45.9	1095	4	AUJ33357	AuJ33357 Human bre	
16	2372.5	45.9	1095	5	ABG78924	Abg78924 Human bre	
17	2372.5	45.9	1095	6	ABU37747	Abu37747 Human tum	
18	2372.5	45.9	1095	7	ADL93315	Adl93315 Human bre	
19	2367	45.8	1002	4	AUJ33351	AuJ33351 Human bre	
20	2367	45.8	1002	5	ABG78918	Abg78918 Human bre	
21	2367	45.8	1002	6	ABU37741	Abu37741 Human tum	
22	2367	45.8	1002	7	ADL93317	Adl93317 Human bre	
23	2367	45.8	1002	8	ADE84427	Ade84427 Human bre	
24	1758	34.0	661	6	ABU37782	Abu37782 Human tum	
25	1758	34.0	661	7	ADL93214	Adl93214 Human bre	

Query Match		100.0%;	Score 5173;	DB 4;	Length 1011;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1011; Conservative		0;	Mismatches	0;	Gaps 0;
QY	1	1	1	1	1
DB	1	1	1	1	1
QY	61	1	1	1	1
DB	61	1	1	1	1
QY	121	1	1	1	1
DB	121	1	1	1	1
QY	181	1	1	1	1
DB	181	1	1	1	1
QY	241	1	1	1	1
DB	241	1	1	1	1
QY	301	1	1	1	1
DB	301	1	1	1	1
QY	361	1	1	1	1
DB	361	1	1	1	1
QY	421	1	1	1	1
DB	421	1	1	1	1
QY	481	1	1	1	1
DB	481	1	1	1	1
QY	541	1	1	1	1
DB	541	1	1	1	1
QY	601	1	1	1	1
DB	601	1	1	1	1
QY	661	1	1	1	1
DB	661	1	1	1	1
QY	721	1	1	1	1
DB	721	1	1	1	1
QY	781	1	1	1	1
DB	781	1	1	1	1
QY	841	1	1	1	1
DB	841	1	1	1	1
QY	901	1	1	1	1
DB	901	1	1	1	1
QY	961	1	1	1	1
DB	961	1	1	1	1

RESULT 2

ADJ69751	ADU69751 standard; protein; 1011 AA.
XX	ADJ69751;
AC	06-MAY-2004 (first entry)
DT	Human heat mitochondrial protein as a therapeutic target SeqID1557.
DE	mitochondrial; human; screening assay; diabetes mellitus;
XX	Huntington's disease; osteoarthritis;
XX	Leber's hereditary optic neuropathy; LHON;
KW	mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW	myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW	neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;
KW	osteopathic; ophthalmological; cytostatic.
OS	Homo sapiens.
XX	WO2003087768-A2.
PN	23-OCT-2003.
XX	04-APR-2003; 2003WO-US010870.
PF	12-APR-2002; 2002US-0372843P.
XX	17-JUN-2002; 2002US-038987P.
PR	20-SEP-2002; 2002US-0412418P.
XX	(MITO-) MITOKOR.
PA	(BUCK-) BUCK INST AGE RES.
XX	Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GW;
PI	Warnock DE;
XX	WPI; 2003-845369/78.
DR	Identifying a mitochondrial target for drug screening assays and for
XX	treating diseases associated with altered mitochondrial function,
PT	comprises detecting a modified polypeptide in a sample and correlating
PT	with the disease.
PS	Claim 1; SEQ ID NO 1557; 180pp; English.
XX	This invention relates to novel mitochondrial targets that can be used
CC	for therapeutic intervention in treating a disease associated with
CC	altered mitochondrial function. Specifically, it refers to a method for
CC	identifying proteins of the human heart mitochondrial proteome that are
CC	useful for drug screening assays, as well as therapeutic targets. The
CC	present invention describes a method for identifying such proteins that
CC	can be used in the treatment of various diseases associated with altered
CC	mitochondrial function including diabetes mellitus, Huntington's disease,
CC	osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC	encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC	ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC	compositions have neuroprotective, nontropic, antidiabetic,
CC	anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC	cytostatic activities. This polypeptide sequence is a human heart
CC	mitochondrial protein of the invention.
XX	Sequence 1011 AA;
SQ	Query Match 100.0%; Score 5173; DB 7; Length 1011;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 1011; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1
DB	1
QY	61

Db 61 CEGSSEIVGMLLQONVDVFAEDIHGITAERYAAARGVNYHQOLLEHRLKLPKNPONTNP 120
Qy 121 EGTSTCTPDEAALPABRTDPTAESLLEKTPDEAARLVEGTSAKIQCLGKATSGKFEQSTE 180
Db 121 EGTSTCTPDEAALPABRTDPTAESLLEKTPDEAARLVEGTSAKIQCLGKATSGKFEQSTE 180
Qy 181 ETPRKILRPTKETSEKFSWPAKERSRKITWEEKETSVKTECVAGVTPNKTVELEKTSNM 240
Db 181 ETPRKILRPTKETSEKFSWPAKERSRKITWEEKETSVKTECVAGVTPNKTVELEKTSNM 240
Qy 241 IACPKTETSKASTNVDSVSSVEPIFSLFGTRTITENSQCTKVBEEDFNLAATKIIKSAAQNY 300
Db 241 IACPKTETSKASTNVDSVSSVEPIFSLFGTRTITENSQCTKVBEEDFNLAATKIIKSAAQNY 300
Qy 301 TCLPDATYQKDIKTINHIEDQMPPSEKREDEEYSWDSGSLFESSAKTQVCIPESMYQ 360
Db 301 TCLPDATYQKDIKTINHIEDQMPPSEKREDEEYSWDSGSLFESSAKTQVCIPESMYQ 360
Qy 361 KWEINREVEELPEKPSAFKPAVEMQKTPNKAFAELKNEQTILRAAQMFPSKQKDEEN 420
Db 361 KWEINREVEELPEKPSAFKPAVEMQKTPNKAFAELKNEQTILRAAQMFPSKQKDEEN 420
Qy 421 SWDESSEPCETVSQKDVYLPKATHQKEFDTLSGKLESPVKDGLLKPCTGCRKVSPLNKALE 480
Db 421 SWDESSEPCETVSQKDVYLPKATHQKEFDTLSGKLESPVKDGLLKPCTGCRKVSPLNKALE 480
Qy 481 LKDRFTFAESPDKDGLLKPCTGCRKVSPLNKALELKDRETLKAESPNDGDLKPTCGRKV 540
Db 481 LKDRFTFAESPDKDGLLKPCTGCRKVSPLNKALELKDRETLKAESPNDGDLKPTCGRKV 540
Qy 541 SLPNKALELKDRETFKAAQMFPSKQKDEENSWDFESFLETLLQNDVCLPKATHQKEF 600
Db 541 SLPNKALELKDRETFKAAQMFPSKQKDEENSWDFESFLETLLQNDVCLPKATHQKEF 600
Qy 601 DTLSGKLESPDKDGLLKPCTGCMKLSLPNKALELKDRETFKAEDEVSSVESTSLGKPTT 660
Db 601 DTLSGKLESPDKDGLLKPCTGCMKLSLPNKALELKDRETFKAEDEVSSVESTSLGKPTT 660
Qy 661 ENSQSTKVEEDFNLTITKEGATKTVTQQRDGIITIERAPQDQTNKMPSELGKEDTKST 720
Db 661 ENSQSTKVEEDFNLTITKEGATKTVTQQRDGIITIERAPQDQTNKMPSELGKEDTKST 720
Qy 721 SDSEIISVSQDTONYECLEPQATYQKEIKTTNGKIESPEKPSHPEPATENQNSVFNKGLEW 780
Db 721 SDSEIISVSQDTONYECLEPQATYQKEIKTTNGKIESPEKPSHPEPATENQNSVFNKGLEW 780
Qy 781 KXKQTLRADSTTLTKILDALPSCERGLKNDKNCQITAKMEQKNNKFCVLOKELSEAKE 840
Db 781 KXKQTLRADSTTLTKILDALPSCERGLKNDKNCQITAKMEQKNNKFCVLOKELSEAKE 840
Qy 841 IKSQLENQKAKWEQELCSVRLPLNQBEEKRRNVVDILKEKIRPEEQRLKLEVKHOLEQTL 900
Db 841 IKSQLENQKAKWEQELCSVRLPLNQBEEKRRNVVDILKEKIRPEEQRLKLEVKHOLEQTL 900
Qy 901 RIQDIELKSVTNLQNVSHSENDLPHENCMWKKEIAMLKLEVATLKHOHVKENKYF 960
Db 901 RIQDIELKSVTNLQNVSHSENDLPHENCMWKKEIAMLKLEVATLKHOHVKENKYF 960
Qy 961 EDIKILQEKNAELQMTLKLKQKTVTKRASQYREQLKVLTAENTMLTSKLKE 1011
Db 961 EDIKILQEKNAELQMTLKLKQKTVTKRASQYREQLKVLTAENTMLTSKLKE 1011

RESULT 3

ABJ37789
ID ABJ37789 standard; protein; 1239 AA.

XX AC ABJ37789;

XX DT 15-MAY-2003 (first entry)

XX DE Human tumour-related protein - SEQ ID No 577.

XX

KW Human; vaccine; gene therapy; T cell stimulation; T cell expansion;
KW tumour; breast cancer; cancer; immune response stimulation.
OS Homo sapiens.
PN WO200283956-A1.
XX 24-OCT-2002
XX 15-APR-2002; 2002WO-US012378.
XX 13-APR-2001; 2001US-00834759.
PR 07-DEC-2001; 2001US-00007805.
PR 13-FEB-2002; 2002US-00076622.
XX (CORI-) CORIXA CORP.
XX Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;
PI Vedwick TS, McNeill PD, Durham M;
XX WPI; 2003-103376/09.
XX New polypeptide and polynucleotide useful for stimulating and/or
PT expanding T cells specific for a tumor protein and treating breast
PT cancer.
XX Example 13; Page 359-363; 375pp; English.
XX The invention comprises a method of stimulating and/or expanding T cells
CC specific for a tumour protein. The invention further comprises human
CC nucleic acids and proteins that are associated with tumours (e.g. breast
CC cancer). The method and sequences of the invention are useful for
CC stimulating and/or expanding T cells specific for a tumour protein,
CC detecting the presence of cancer, stimulating an immune response in a
CC patient and treating breast cancer. The present amino acid sequence
CC represents a human tumour-related protein
XX
SQ Sequence 1239 AA;

Query Match 100.0%; Score 5173; DB 6; Length 1239;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1011; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVATLLSYGAVIEVQNKASLTPLLLAIQKRSQTVFELLTKNANAFNESCKTALMLAI 60
Db 1 MVATLLSYGAVIEVQNKASLTPLLLAIQKRSQTVFELLTKNANAFNESCKTALMLAI 60
Qy 61 CEGSSEIVGMLLQONVDVFAEDIHGITAERYAAARGVNYHQOLLEHRLKLPKNPONTNP 120
Db 61 CEGSSEIVGMLLQONVDVFAEDIHGITAERYAAARGVNYHQOLLEHRLKLPKNPONTNP 120
Qy 121 EGTSTCTPDEAALPABRTDPTAESLLEKTPDEAARLVEGTSAKIQCLGKATSGKFEQSTE 180
Db 121 EGTSTCTPDEAALPABRTDPTAESLLEKTPDEAARLVEGTSAKIQCLGKATSGKFEQSTE 180
Qy 181 ETPRKILRPTKETSEKFSWPAKERSRKITWEEKETSVKTECVAGVTPNKTVELEKTSNM 240
Db 181 ETPRKILRPTKETSEKFSWPAKERSRKITWEEKETSVKTECVAGVTPNKTVELEKTSNM 240
Qy 241 IACPKTETSKASTNVDSVSSVEPIFSLFGTRTITENSQCTKVBEEDFNLAATKIIKSAAQNY 300
Db 241 IACPKTETSKASTNVDSVSSVEPIFSLFGTRTITENSQCTKVBEEDFNLAATKIIKSAAQNY 300
Qy 301 TCLPDATYQKDIKTINHIEDQMPPSEKREDEEYSWDSGSLFESSAKTQVCIPESMYQ 360
Db 301 TCLPDATYQKDIKTINHIEDQMPPSEKREDEEYSWDSGSLFESSAKTQVCIPESMYQ 360
Qy 361 KWEINREVEELPEKPSAFKPAVEMQKTPNKAFAELKNEQTILRAAQMFPSKQKDEEN 420
Db 361 KWEINREVEELPEKPSAFKPAVEMQKTPNKAFAELKNEQTILRAAQMFPSKQKDEEN 420
Qy 421 SWDESSEPCETVSQKDVYLPKATHQKEFDTLSGKLESPVKDGLLKPCTGCRKVSPLNKALE 480

Db 421 SWDSSEPCETVSQKQVYLPKATHQKEFDLSGLKEESPVKDGLLKTCTGGRKVSPLNKALE 480
Qy 481 LKDRFTFAESPDKDGLLKTCTGGRKVSPLNKALELKDRFTLKAESPNDGLLKTCTGGRKV 540
Db 481 LKDRFTFAESPDKDGLLKTCTGGRKVSPLNKALELKDRFTLKAESPNDGLLKTCTGGRKV 540
Qy 541 SLPNKALELKDRFTFAESPDKDGLLKTCTGGRKVSPLNKALELKDRFTLKAESPNDGLLKTCTGGRKV 600
Db 541 SLPNKALELKDRFTFAESPDKDGLLKTCTGGRKVSPLNKALELKDRFTLKAESPNDGLLKTCTGGRKV 600
Qy 601 DTLGKLEESPDKDGLLKTCTGGRKVSPLNKALELKDRFTLKAESPNDGLLKTCTGGRKV 660
Db 601 DTLGKLEESPDKDGLLKTCTGGRKVSPLNKALELKDRFTLKAESPNDGLLKTCTGGRKV 660
Qy 661 ENSQSTKVEEDFNLTATKEGATKTVTGQQRDGIITIERAPQDQTNKMPSELGRKEDTKST 720
Db 661 ENSQSTKVEEDFNLTATKEGATKTVTGQQRDGIITIERAPQDQTNKMPSELGRKEDTKST 720
Qy 721 SDSEIISVSQDQNYECLPEATYQKEIKTTNGKIEESPXKPSHFEPATEMONSVNKGLEW 780
Db 721 SDSEIISVSQDQNYECLPEATYQKEIKTTNGKIEESPXKPSHFEPATEMONSVNKGLEW 780
Qy 781 KKKOTLRADSTTLKLDALPSCERGRELKONCEQITAKMEQMKNFVLOKELSEAKE 840
Db 781 KKKOTLRADSTTLKLDALPSCERGRELKONCEQITAKMEQMKNFVLOKELSEAKE 840
Qy 841 IKSQLEQKAKWEQELCSVRLPLNQBEKRRNVDIILKEKIRPEEQLRKKLEVKHQLEQTL 900
Db 841 IKSQLEQKAKWEQELCSVRLPLNQBEKRRNVDIILKEKIRPEEQLRKKLEVKHQLEQTL 900
Qy 901 RIQDIELKSVTSNLQVSHTHSENDLFHENCMLKKEIAMLKLEVATLKHQOVKENYF 960
Db 901 RIQDIELKSVTSNLQVSHTHSENDLFHENCMLKKEIAMLKLEVATLKHQOVKENYF 960
Qy 961 EDIKILQEKNAELQMTLKLKQKTVTKRASQYRQLKVLTAENTMLTSKLKE 1011
Db 961 EDIKILQEKNAELQMTLKLKQKTVTKRASQYRQLKVLTAENTMLTSKLKE 1011

RESULT 4
ADL93239
ID ADL93239 standard; protein; 1239 AA.
XX AC ADL93239;
XX DT 20-MAY-2004 (first entry)
XX DE Human breast cancer-associated polypeptide #54.
XX KW gene therapy; protein therapy; vaccine; breast cancer; cancer; human.
XX OS Homo sapiens.
XX PN US2003166022-A1.
XX PD 04-SEP-2003.
XX PF 15-APR-2002; 2002US-00124805.
XX PR 28-DEC-1998; 98US-00222575.
XX PR 02-APR-1999; 99US-00285480.
XX PR 23-JUN-1999; 99US-00339338.
XX PR 02-SEP-1999; 99US-00389681.
XX PR 03-NOV-1999; 99US-00433826.
XX PR 17-APR-2000; 2000US-00551621.
XX PR 08-JUN-2000; 2000US-00590751.
XX PR 22-JUN-2000; 2000US-00604287.
XX PR 20-JUL-2000; 2000US-00620405.
XX PR 13-APR-2001; 2001US-00834759.
XX PR 07-DEC-2001; 2001US-00007805.
XX PR 13-FEB-2002; 2002US-00076622.

(CORI-) CORIXA CORP.
Houghton RL, Sleath PR, Persing DH;
WPI; 2003-874918/81.
N-PSDB; ADL93238.
An isolated oncogenic polypeptide useful for preventing, diagnosing and treating breast cancer.
Example 13; SEQ ID NO 577; 294pp; English.
The invention relates to an isolated breast cancer-associated polypeptide. The polypeptide may be used for the diagnosis and treatment of breast cancers. The methods are useful for detecting the presence of a cancer in a patient and treating a cancer in a patient. The present sequence represents the amino acid sequence of a human breast cancer-associated polypeptide.
Sequence 1239 AA;
Query Match 100.0%; Score 5173; DB 7; Length 1239;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1011; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MVATLLSYCAVIEVQNKASLTPLLLAIQKRSKQTVFELTKNANANAFNEKCTALMLAI 60
Db 1 MVATLLSYCAVIEVQNKASLTPLLLAIQKRSKQTVFELTKNANANAFNEKCTALMLAI 60
Qy 61 CEGSSEIVGMLLQNVVPAEDIHGIITAERYAARGVNYIHQOLLEHIRKLPKNQNTNP 120
Db 61 CEGSSEIVGMLLQNVVPAEDIHGIITAERYAARGVNYIHQOLLEHIRKLPKNQNTNP 120
Qy 121 EGTSTGTDPDEAPLAERTPDTAESLEKTPDEAARLVGTSAKIOCLGKATSGKEQSTE 180
Db 121 EGTSTGTDPDEAPLAERTPDTAESLEKTPDEAARLVGTSAKIOCLGKATSGKEQSTE 180
Qy 181 ETPRKILRPTKETSEKFWPAKERSKJTWEEKTSVKTECVAGVTPNKTEVLEKGTSM 240
Db 181 ETPRKILRPTKETSEKFWPAKERSKJTWEEKTSVKTECVAGVTPNKTEVLEKGTSM 240
Qy 241 IACPTKETSTKASTNVDSVSEPIFSLFGTRIENSQCTKVEEDFNLATKIISKAAQNY 300
Db 241 IACPTKETSTKASTNVDSVSEPIFSLFGTRIENSQCTKVEEDFNLATKIISKAAQNY 300
Qy 301 TCLPDATYQKDIKTINHIEDQMPFSESKREDEEYSWDSGLSFESSAKTQVCIPESMYQ 360
Db 301 TCLPDATYQKDIKTINHIEDQMPFSESKREDEEYSWDSGLSFESSAKTQVCIPESMYQ 360
Qy 361 KYWEINREVEELPEKPSAFKPAVEMQKTVPNKAFELKNEQTLRAAQMPFSESKQKDDEN 420
Db 361 KYWEINREVEELPEKPSAFKPAVEMQKTVPNKAFELKNEQTLRAAQMPFSESKQKDDEN 420
Qy 421 SWDSSEPCETVSQKQVYLPKATHQKEFDLSGLKEESPVKDGLLKTCTGGRKVSPLNKALE 480
Db 421 SWDSSEPCETVSQKQVYLPKATHQKEFDLSGLKEESPVKDGLLKTCTGGRKVSPLNKALE 480
Qy 481 LKDRFTFAESPDKDGLLKTCTGGRKVSPLNKALELKDRFTLKAESPNDGLLKTCTGGRKV 540
Db 481 LKDRFTFAESPDKDGLLKTCTGGRKVSPLNKALELKDRFTLKAESPNDGLLKTCTGGRKV 540
Qy 541 SLPNKALELKDRFTFAESPDKDGLLKTCTGGRKVSPLNKALELKDRFTLKAESPNDGLLKTCTGGRKV 600
Db 541 SLPNKALELKDRFTFAESPDKDGLLKTCTGGRKVSPLNKALELKDRFTLKAESPNDGLLKTCTGGRKV 600
Qy 601 DTLGKLEESPDKDGLLKTCTGGRKVSPLNKALELKDRFTLKAESPNDGLLKTCTGGRKV 660
Db 601 DTLGKLEESPDKDGLLKTCTGGRKVSPLNKALELKDRFTLKAESPNDGLLKTCTGGRKV 660
Qy 661 ENSQSTKVEEDFNLTATKEGATKTVTGQQRDGIITIERAPQDQTNKMPSELGRKEDTKST 720
Db 661 ENSQSTKVEEDFNLTATKEGATKTVTGQQRDGIITIERAPQDQTNKMPSELGRKEDTKST 720

QY 721 SDSEIISVSDTQNYECLPEATYQKEIKTTNGKIBESPEKPSHPFEPATEMNSVFNKGLW 780
 Db 721 SDSEIISVSDTQNYECLPEATYQKEIKTTNGKIBESPEKPSHPFEPATEMNSVFNKGLW 780
 QY 781 KKKQTLRADSTTLTKILDALPSCERGLKKDNCQITAKMEQMKNFVQLKELSEAKE 840
 Db 781 KKKQTLRADSTTLTKILDALPSCERGLKKDNCQITAKMEQMKNFVQLKELSEAKE 840
 QY 841 KKSQLENQAKWEQELCSVRLPLNQEEKKRRNVDILKEKIRPEEQLRKKLEVKHOLEQTL 900
 Db 841 KKSQLENQAKWEQELCSVRLPLNQEEKKRRNVDILKEKIRPEEQLRKKLEVKHOLEQTL 900
 QY 901 RIQDIELKSVTSNLNQVSHTHSENDLPHENCMLKKEIAMLKLEVATLKHQHVKNKYF 960
 Db 901 RIQDIELKSVTSNLNQVSHTHSENDLPHENCMLKKEIAMLKLEVATLKHQHVKNKYF 960
 QY 961 EDIKILOKNAELQMTLKLKQKTVTKRASQYREQLKVLTAENTMLTSKLKE 1011
 Db 961 EDIKILOKNAELQMTLKLKQKTVTKRASQYREQLKVLTAENTMLTSKLKE 1011

RESULT 5

ABR47547
 ID ABR47547 standard; protein; 1225 AA.

AC ABR47547;

DT 12-JUN-2003 (first entry)

DE Breast cancer associated protein sequence SEQ ID NO:332.

KW Human; breast cancer; cytostatic; gene therapy.

XX Homo sapiens.

XX WO2003004989-A2.

XX 16-JAN-2003.

XX 21-JUN-2002; 2002WO-US019669.

XX 21-JUN-2001; 2001US-0299887P.

PR 27-JUN-2001; 2001US-0301572P.

PR 18-JUL-2001; 2001US-0306501P.

PR 25-SEP-2001; 2001US-0325002P.

PR 05-MAR-2002; 2002US-0362585P.

PR 14-MAY-2002; 2002US-0380391P.

XX (MILL-) MILLENIUM PHARM INC.

XX Lillie J, Gannavarapu M, Glatt K, Hoersh S, Kamatkar S;

PI Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;

PI Bast RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;

XX WPI; 2003-210381/20.

DR N-ESDB; ACC50245.

PT Breast cancer diagnosis or treatment by comparing the level of expression

PT of a marker in a patient sample with that in the control non-breast

XX cancer sample.

XX Claim 1; SEQ ID NO 332; 128pp; English.

XX The present invention describes a method for assessing whether a patient

CC is afflicted with breast cancer. The method comprises comparing the level

CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and

CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1225 AA;

Query Match 99.7%; Score 5156; DB 6; Length 1225;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1008; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MVATLLSGAVIEQNKASLTPLLLAIQKRSKQTVFELLTKNANANAFNESCCTALMLAI 60
 Db 154 MVATLLSGAVIEQNKASLTPLLLAIQKRSKQTVFELLTKNANANAFNESCCTALMLAI 213
 QY 61 CEGSEIIGMLLQONVDVPAEDIHGITAERVAARGVNYIHQOLLEHRLKLPKNQNTNP 120
 Db 214 CEGSEIIGMLLQONVDVPAEDIHGITAERVAARGVNYIHQOLLEHRLKLPKNQNTNP 273
 QY 121 EGTSTGTPDEAAPLAERTPDTAESLLEKTPDEAARLVGTSAKIQCLGKATSGKFEQSTE 180
 Db 274 EGTSTGTPDEAAPLAERTPDTAESLLEKTPDEAARLVGTSAKIQCLGKATSGKFEQSTE 333
 QY 181 ETPRKILRPYKTSKFSWPAKERSRKITWEEKTSVKTECVAGVTPNKTVELEKGTSM 240
 Db 334 ETPRKILRPYKTSKFSWPAKERSRKITWEEKTSVKTECVAGVTPNKTVELEKGTSM 393
 QY 241 IACPTKETSTKASTNVDSVVEPIPSLFGTITENSQCTKVEEDFNLATKIISKSAANY 300
 Db 394 IACPTKETSTKASTNVDSVVEPIPSLFGTITENSQCTKVEEDFNLATKIISKSAANY 453
 QY 301 TCLPDATYQKDIKTINHKKIEDQMPFSESKREDEBEYSWDSGSLFESSAKTQVCIPESMYQ 360
 Db 454 TCLPDATYQKDIKTINHKKIEDQMPFSESKREDEBEYSWDSGSLFESSAKTQVCIPESMYQ 513
 QY 361 KWMEINREVEELPEKPSAFKPAVEQKTPNKAELKNEQTLRAAQMPFSESKQKDDEN 420
 Db 514 KWMEINREVEELPEKPSAFKPAVEQKTPNKAELKNEQTLRAAQMPFSESKQKDDEN 573
 QY 421 SWDSSEPCETVSQKDVYLPKATHQKEFDTLGKLEESPVKDGLLKPTCGRKVSUKNKALE 480
 Db 574 SWDSSEPCETVSQKDVYLPKATHQKEFDTLGKLEESPVKDGLLKPTCGRKVSUKNKALE 633
 QY 481 LKDRETFKAESPDKDGLLKPTCGRKVSUKNKALEKDRETLKAESPDKDGLLKPTCGRKRV 540
 Db 634 LKDRETFKAESPDKDGLLKPTCGRKVSUKNKALEKDRETLKAESPDKDGLLKPTCGRKRV 693
 QY 541 SLPNKALELKDRETFKAQAMPFSESKQKDDENSWDFESFLETLQNDVCLPKATHQKEF 600
 Db 694 SLPNKALELKDRETFKAQAMPFSESKQKDDENSWDFESFLETLQNDVCLPKATHQKEF 753
 QY 601 DTLGKLEESPDKDGLLKPTCGMKISLPNKALELKDRETFKAEDVSSVSESTFSLFGKPTT 660
 Db 754 DTLGKLEESPDKDGLLKPTCGMKISLPNKALELKDRETFKAEDVSSVSESTFSLFGKPTT 813
 QY 661 ENSQSTKVEEDFNLTITKEGATKTVTGQERDIGIIERAPQDQTNKMPSTSELGRKEDTKST 720
 Db 814 ENSQSTKVEEDFNLTITKEGATKTVTGQERDIGIIERAPQDQTNKMPSTSELGRKEDTKST 873
 QY 721 SDSIIISVSDTQNYECLPEATYQKEIKTTNGKIBESPEKPSHPFEPATEMNSVFNKGLW 780
 Db 874 SDSIIISVSDTQNYECLPEATYQKEIKTTNGKIBESPEKPSHPFEPATEMNSVFNKGLW 933
 QY 781 KKKQTLRADSTTLTKILDALPSCERGLKKDNCQITAKMEQMKNFVQLKELSEAKE 840
 Db 934 KKKQTLRADSTTLTKILDALPSCERGLKKDNCQITAKMEQMKNFVQLKELSEAKE 993
 QY 841 KKSQLENQAKWEQELCSVRLPLNQEEKKRRNVDILKEKIRPEEQLRKKLEVKHOLEQTL 900
 Db 994 KKSQLENQAKWEQELCSVRLPLNQEEKKRRNVDILKEKIRPEEQLRKKLEVKHOLEQTL 1053
 QY 901 RIQDIELKSVTSNLNQVSHTHSENDLPHENCMLKKEIAMLKLEVATLKHQHVKNKYF 960
 Db 1054 RIQDIELKSVTSNLNQVSHTHSENDLPHENCMLKKEIAMLKLEVATLKHQHVKNKYF 1113

Tue Feb 22 09:49:15 2005

QY	961	EDIKILOEKNAELQMTLTKQKQVTKRASQYRQOLKVLTAENTMLTSKKE	1011
Db	1114	EDIKILOEKNAELQMTLTKQKQVTKRASQYRQOLKVLTAENTMLTSKKE	1164
RESULT 6			
ID	AA	AAB84702	
XX	AA	AAB84702 standard; protein; 1341 AA.	
AC	AA	AAB84702;	
XX	17-SEP-2001	(first entry)	
DE	Amino acid sequence of a human cancer associated antigen.		
XX	Cancer associated antigen; INGI; tumour suppressor; cancer; vaccine.		
OS	Homo sapiens.		
XX	WO200147959-A2.		
XX	05-JUL-2001.		
XX	29-NOV-2000; 2000WO-US042334.		
XX	30-NOV-1999; 99US-00451739.		
XX	24-OCT-2000; 2000US-00602362.		
XX	(LUDW-) LUDWIG INST CANCER RES.		
XX	(SLOK) SLOAN KETTERING INST CANCER RES.		
XX	(CORR) CORNELL RES FOUND INC.		
XX	Jager D, Stockert E, Scanlan M, Knuth A, Old L, Gure A, Chen Y;		
XX	WPI; 2001-441706/47.		
XX	Isolated cancer associated nucleic acid molecule identified by SEREX		
XX	(serological identification of antigens by recombinant expression		
XX	cloning) technique, useful in nucleic acid based therapies to treat		
XX	cancer.		
XX	Claim 83; Page 53-57; 62pp; English.		
XX	The present sequence represents a human cancer associated antigen. The		
XX	sequence was identified using probes derived from the INGI gene. The INGI		
XX	gene is a tumour suppressor candidate gene. The cancer associated antigen		
XX	polynucleotides and polypeptides are useful for screening for the		
XX	possible presence of a pathological condition in a subject such as		
XX	cancer. The cancer associated antigen polypeptides are useful for		
XX	producing vaccines		
SQ	Sequence 1341 AA;		
Query Match			
Best Local Similarity 59.4%; Score 3075; DB 4; Length 1341;			
Matches 664; Conservative 57.4%; Pred. No. 3.8e-199;			
1 MYATLLSGAVIEVQNKASLPLLLAIOKRSKQTVFELLTKNANAFNESCCTALMLAI 60			
98 VVAKLLSHGAVIEVHNKASLPLLLSITRSEQIVFELLKNANAVNKYKCTALMLAV 157			
61 CEGSSEIVGMLQQNVDFAEIHGITAERYAARGVNYIHQOLLEHRLKLPKQNTNP 120			
158 CEGSSEIVGMLQQNVDFAAIDICGVTAHYAVTCGFHHIHEQIMYIRKLSKKHNTNP 217			
121 EGTSTGTPDEAAPLA-----ERTPDATSLLEKTPDEAARLVE 158			
218 EGTSTGTPDEAAPLAERTPDATSLVEKTPDEAAPLVKTPDAESLVEKTPDEAASLVE 277			
159 GTSAKIOCLCKATSGFEGOSTETPRKILRPKTETSEKFSWP----- 200			
278 GTSDKIOCLEKATSGFEGSAETPREITSPAKETSEKFTWPAKGRPKIAWEKKETDTPR			

QY	201	-----AKERSRKITWEEKETSVKTECVAGVTTPNKTEVLEKGTSNIAACP	244
Db	338	EIMSPAKETSEKFTWPAKGRPKIAWEKKETPVKTGCVARVTSNKTVKLEGRSRMIAACP	397
QY	245	TKETSTKASTNVDVSSVEPIFLSGTRTITENSQCTKVBEEDFNLAITKISKAQNYTCLP	304
Db	398	TKESSTKASAN-----	408
QY	305	DATYQKOIKTINHIEDOMPSESKEDEEYSWDSGSLFESSAKTQVCIPESMYQKME	364
Db	409	-----DQFPSESKEDEEYSCDSRSLFESSAKIQVCIPESIYQKME	452
QY	365	INREVEELPEKPSAFKPAVEMQKTPNKAFAELKNEQTLRAAQMPFSESQKQDEENSWS	424
Db	453	INREVEEPPKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADPMPFPFSESQKQDEENSWS	512
QY	425	ESPCETVSQKDVYLPKATHQKEFTLSGKLEESPVKDGLLKPTCGRKVSLPNKALELKOR	484
Db	513	ESLCETVSQKDVCLPKATHQKEIDKINGKLEESPVKDGLLKATCGMKVSIPTKALELKD	572
QY	485	ETFKAESPKDKGLLKPTCGRKVSLPNKALELKORETLKAESPDNDGLLKPTCGRKVSLPN	544
Db	573	QTFKAEPPGKPSAFEP-----ATEMQK-----SVPN	598
QY	545	KALELKDRETFKAQMPFSESQKQDEENSWSDFLETLQNDVCLPKATHQKEFTLS	604
Db	599	KALELKNEQTLRADPMPFPFSESQKQDEENSWSDFLETLQNDVCLPKATHQKEIDKIN	658
QY	605	GKLEESPDKDGLLKPTCGMKISLPNKALELKDRETFKAEDVSSVSEFTSLFGKPTTENSQ	664
Db	659	GKLEGSPVKDGLLKANCCKMKSIPITKALELMDQMTFFKAEPEK-PSAF-----EPAIEMQK	713
QY	665	STKVEEDFNLTKEGATKTVTQGBRDIGIIERAQDQTNKMTSELGKEDTKSTDSSE	724
Db	714	SV-----PNKALELKNEQTLRADPILP-----SESKQKDYEESSWDSSE	751
QY	725	IISVSDTONYECPLPEATYQKEIKTTNGKLEES-----	756
Db	752	SLCETVSQKDVCLPKATHQKEIDKINGKLEESPNDNDGFLKAPCRMKVSIPTKALELMDQ	811
QY	757	-----PEKPSHPEPATEMQNSVPNKGLEWKNKQTLRA-----	788
Db	812	TFKAEPPKPSAFKPAIEMQKSVNKALELKNEQTLRADQMPFSESQKQVEENSWSDES	871
QY	789	-----DSTLSKILDALPSCERGRKELKKDNCEQITA	819
Db	872	LRETVSQKDVCPKATHQKEMDKISGKLEBDSLSKILDTVISCERARELOKDHCEQRTG	931
QY	820	KMEQMKKFCVLOKELSEAEIKSQLENOKAKWEQELCSVRLPQNEEKKRNVDILKEK	879
Db	932	KMEQMKKFCVLLKKLSEAEIKSQLENOKAKWEQELCSVRLTLNQEERKRNADILNEK	991
QY	880	IRP-----EEQURKXLEVKGHLEQTLRIQDIELKSVTSNLNOVSHTHSENDLFHENCML	934
Db	992	IREEGLRIEEQHRKELEVQKQLEQALRIODIELKSVESNLNOVSHTHSENYLLHENCML	1051
QY	935	KKEIAMLKLEVATLKHQHOVKENKVFEDIKLOEKNAELQMTLTKQKQVTKRASQYREQ	994
Db	1052	KKEIAMLKLEIATLKHQHOVEKENKVFEDIKILKEKNAELQMTLTKKXESLTGRASQYSGQ	1111
QY	995	LKVLTAENTMLTSKLKE 1011	
Db	1112	LKVLTAENTMLTSKLKE 1128	
RESULT 7			
ID	AB	ABJ05537	
XX	AB	ABJ05537 standard; protein; 1341 AA.	
AC	AB	ABJ05537;	
DT	14-NOV-2002	(first entry)	

RESULT 7
ABJ05537
ID ABJ05537 standard; protein; 1341 AA.
XX
AC ABJ05537;
XX
DT 14-NOV-2002 (first entry)

XX	Breast cancer-associated protein 2.		
XX	Breast cancer; breast cancer-associated gene sequence; drug development;		
KW	pharmacogenetics; biosensor development.		
XX	Unidentified.		
OS			
XX	WO200259377-A2.		
PN			
XX	01-AUG-2002.		
XX			
PF	24-JAN-2002; 2002WO-US002242.		
XX			
PR	24-JAN-2001; 2001US-0263965P.		
PR	02-FEB-2001; 2001US-0265928P.		
PR	09-APR-2001; 2001US-00829472.		
PR	09-APR-2001; 2001US-0283698P.		
PR	04-MAY-2001; 2001US-0288590P.		
PR	29-MAY-2001; 2001US-0294443P.		
XX			
PA	(EOSB-) EOS BIOTECHNOLOGY INC.		
XX			
XX	Mack DH, Gish KC, Afar D;		
PI			
XX			
DR	WPI; 2002-583738/62.		
DR	N-PSDB; ABT07694.		
XX			
PT	Detecting a breast cancer-associated transcript in a patient's cell,		
PT	useful for diagnosing breast cancer, comprises contacting a biological		
PT	sample with a polynucleotide that selectively hybridizes with breast		
PT	cancer nucleic acids.		
XX			
PS	Disclosure; Page 348-349; 414pp; English.		
XX			
CC	The invention comprises a method of detecting a breast cancer-associated		
CC	transcript in a cell from a patient. The method of the invention involves		
CC	contacting a biological sample from the patient with a nucleotide that		
CC	hybridizes to one of the 69 breast cancer-associated gene sequences shown		
CC	in the specification. The method of the invention is useful in the		
CC	diagnosis or prognosis of breast cancer, and for detecting genes that are		
CC	up or down-regulated in breast cancer cells. Genes identified by the		
CC	method of the invention can be used in diagnostic purposes and also as		
CC	targets for screening for therapeutic compounds that modulate breast		
CC	cancer (e.g. hormones or antibodies). Identification of genes that are		
CC	over or under expressed in breast cancer can additionally provide high-		
CC	resolution, high-sensitivity datasets which can be used in the areas of		
CC	diagnostics, therapeutics, drug development, pharmacogenetics, protein		
CC	structure and biosensor development. Amino acid sequences ABJ05536 -		
CC	ABJ05604 represent the proteins encoded by the 69 breast cancer-		
CC	associated genes of the invention		
XX			
SQ	Sequence 1341 AA;		
	Query Match 59.4%; Score 3075; DB 5; Length 1341;		
	Best Local Similarity 57.4%; Pred. No. 3.8e-199;		
	Matches 664; Conservative 76; Mismatches 145; Indels 272; Gaps 12;		
QY	1 MVATLLSGAVTEVQNKASLTPLLLAIQKRSQTVEFLITKANANAFNESKCTALMLAI 60		
DB	98 VVAKLLSHGAVTEVHNKASLTPLLLSITKRSQIIVEFLLIKANANAVNKYCTALMLAV 157		
QY	61 CEGSSEIVGMLLQQNVDFVFAEDIHGITAERYAAARGVNYIHOOLLEHRIKLPKNQNTNP 120		
DB	158 CHGSSEIVGMLLQQNVDFVFAEDICGVTAEHYAVTCGFHHIEQIWEYIRKLSKNQNTNP 217		
QY	121 EGTSTGTDPDEAAPLA-----ERTPTAESLLEKTPDEAARLVE 158		
DB	218 ECTSAGTPDEAAPLAERTPDTAESLVEKTPDEAAPLVERTPDTAESLVEKTPDEAASLVE 277		
QY	159 GTSIAKIQCLGKATSGKFGQSTETETPRKILRPKETSFKFSWP----- 200		
DB	278 GTSDKIQCLEKATSGKFGQSAETPREITSPAKETSEKFTWPAKGRPKIAWEKKEDTPR 337		

RESULT 8

ABJ37784

ID ABJ37784 standard; protein; 1341 AA.

XX ABJ37784;

XX DT 15-MAY-2003 (first entry)

409 -----DQRPSPSEKQDEDEEYSCDSRSLFESSAKIQVICIPESIIYQKWE 452
365 INREVELPEKPSAPKPAVEMQKTPVNFKAFLKNEQTLRAAOMFPSESQKQDEENSWDS 424
453 INREVEBPKPSPAFKPAJEMQNSVPNKAPELKNQTLRADPMFPPEKQKQDEENSWDS 512
425 EPCETVSQKDVYLPKATHQKBFDTLSGKLEESPVKDGLLKPTCGRKVSLPKALELKDOR 484
513 ESLCETVSQKDVYLPKATHQKBFDTLSGKLEESPVKDGLLKPTCGRKVSLPKALELKDOR 572
485 ETKFAESPDKGLLKPTCGRKVSLPKALELKDORRETLKAESEPDNDGLLAPTKCGRKVSLPN 544
573 QTFKAPPGKPSAFEP-----ATEMOK-----SVEN 598
545 KALELKDORRETFKAAQMPFPSESQKQDEENSWDSFLETLQNDVCLPKATHQKBFDTLS 604
599 KALELKDORRETFKAAQMPFPSESQKQDEENSWDSFLETLQNDVCLPKATHQKBFDTLS 658
605 GKLEESPDGKLLKPTCGMKISLPNKALELKDORRETFKAEVSVSESTFSLFGPPTTENSQ 664
659 GKLEESPDGKLLKPTCGMKISLPNKALELKDORRETFKAEVSVSESTFSLFGPPTTENSQ 713
665 STKVEEDFNLTTEGATKVTGQOERDGIIRAPQDQTNKMTSELGRKEDTKSTDSE 724
714 SV-----PNKALELKDORRETFKAEVSVSESTFSLFGPPTTENSQ 751
725 IISVSDTQNYECLPEATVQKEIKTTNGKIBES----- 756
752 SLCTVSQKQDVCLPKATHQKBFDTLSGKLEESPVKDGLLKPTCGRKVSLPKALELMDQ 811
757 -----PEKSPHFPATPMQNSVPNKALELKDORRETFKAEVSVSESTFSLFGPPTTENSQ 871
812 TFKAEPPKPSAFKPAJEMQNSVPNKALELKDORRETFKAEVSVSESTFSLFGPPTTENSQ 871
789 -----DSTTLKILDALPSCGERGELKKNCEQITA 819
872 LRETVSQKDVYLPKATHQKBFDTLSGKLEESPVKDGLLKPTCGRKVSLPKALELKDOR 931
820 KMEQMKNFVQLQKELSEAKESQLENQKAKVQELCSVRLPLNOBEKERNVDILKEK 879
932 KMEQMKNFVQLQKELSEAKESQLENQKAKVQELCSVRLPLNOBEKERNVDILKEK 991
880 IRP-----EEQLRKLEVKHOLEQTLRIODIELKSVTSNLNOVSHTHESENDLPHENCM 934
992 IREELGRIEEQHREKLEVKQQLQALRIODIELKSVTSNLNOVSHTHESENDLPHENCM 1051
935 KKEIAMLKLEIATLKHQVQENKYPFEDIKILQKNAELQMTLKLKQKVTGRASQYREQ 994
1052 KKEIAMLKLEIATLKHQVQENKYPFEDIKILQKNAELQMTLKLKQKVTGRASQYREQ 1111
995 LKVLTAENTMLTSKLKE 1011
1112 LKVLTAENTMLTSKLKE 1128
RESULT 9
ABR47548
ID ABR47548 standard; protein; 1341 AA.
XX ABR47548;
AC ABR47548;
XX
DT 12-JUN-2003 (first entry)
XX Breast cancer associated protein sequence SEQ ID NO:334.
DE Human; breast cancer; cytostatic; gene therapy.
XX Homo sapiens.
OS
XX WO2003004989-A2.
FN
XX
PD 16-JAN-2003.
XX

XX Human tumour-related protein - SEQ ID No 565.
DE Human; vaccine; gene therapy; T cell stimulation; T cell expansion;
KW tumour; breast cancer; cancer; immune response stimulation.
XX Homo sapiens.
OS
XX WO200283956-A1.
PN
XX 24-OCT-2002.
XX
PF 15-APR-2002; 2002WO-US012378.
XX
XX 13-APR-2001; 2001US-00834759.
PR 07-DEC-2001; 2001US-00007805.
PR 13-FEB-2002; 2002US-00076622.
XX
XX (CORI-) CORIXA CORP.
PA
XX Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Panger GR;
PI Vedvick TS, McNeill PD, Durham M;
XX
XX WPI; 2003-103376/09.
XX
XX New polypeptide and polynucleotide useful for stimulating and/or
PT expanding T cells specific for a tumor protein and treating breast
PT cancer.
XX
XX Example 9; Page 342-346; 375pp; English.
XX
XX The invention comprises a method of stimulating and/or expanding T cells
CC specific for a tumour protein. The invention further comprises human
CC nucleic acids and proteins that are associated with tumours (e.g. breast
CC cancer). The method and sequences of the invention are useful for
CC stimulating and/or expanding T cells specific for a tumour protein,
CC detecting the presence of cancer, stimulating an immune response in a
CC patient and treating breast cancer. The present amino acid sequence
CC represents a human tumour-related protein
XX
XX Sequence 1341 AA;
Query Match. 59.4%; Score 3075; DB 6; Length 1341;
Best Local Similarity 57.4%; Pred. No. 3.8e-199;
Matches 664; Conservative 76; Mismatches 145; Indels 272; Gaps 12;
1 MVATLLSYGAVIEVQNKASLTPILLATQKRSQTVFELLTKNANANAFNESCCTALMLAI 60
98 VVAKLLSHGAVIEVHNKASLTPILLATQKRSQTVFELLTKNANANAVNKYKCTALMLAV 157
61 CEGSEIVGMLLQNVDFADIEHGITAERYAAAGVNYIHQQLLEHRIKLPGNPQNTNP 120
158 CHGSEIVGMLLQNVDFADIEHGITAERYAAAGVNYIHQQLLEHRIKLPGNPQNTNP 217
121 EGTSTGTGTPDEAPLA-----BRTDPTAESLLEKTPDEAARLVE 158
218 EGTSGTTPDEAPLAERTPTDPAESLVEKTPDEAAPLVEKTPDPAESLVEKTPDEAARLVE 277
159 GTSKIQGLKATSKGFQSTTEETPKILRPTKETSEKFSWP----- 200
278 GTSKIQGLKATSKGFQSTTEETPKILRPTKETSEKFSWP----- 337
201 -----AKERSKITWEKETS SVKTECVAGVTPNKTVELEKGTSMNIACP 244
338 EIMSPAKETSEKFTWAAKGRPKIAWEKETEPTVKTGCVARVTSNKTVELEKGTSMNIACP 397
245 TKETSTKASTNVDSVVEPIFSLFGTRIENSQCTKVEEDFNLATKIISKSAQAQNTCLP 304
398 TKESSTKASAN----- 408
305 DATYQDKINTINHIEDQMPFSGESKREDEEYSWDSGLFESSAKTQVCIPESMYQKWE 364

[illegible]


```
QY 427 PCETVSKDVLKPTKATKQKEDTSLGKLESPVXGDKLLKPTCGKRVSLPNKALELKDRET 486
Db 187 LCETVSKDVLKPTKATKQKEDTSLGKLESPVXGDKLLKPTCGKRVSLPNKALELKDRET 246
QY 487 FKAESPDKDGLLPTKATKQKEDTSLGKLESPVXGDKLLKPTCGKRVSLPNKALELKDRET 546
Db 247 FKAESPDKDGLLPTKATKQKEDTSLGKLESPVXGDKLLKPTCGKRVSLPNKALELKDRET 284
QY 547 LELKDRETFKAAQMPSPSKDDEENSWDPEFLETLQNDVCLPKATHQKEDTSLGK 606
Db 285 -----ELIPSESKDDEENSWDPEFLETLQNDVCLPKATHQKEDTSLGK 332
QY 607 LEESPDKGLLPTKATKQKEDTSLGKLESPVXGDKLLKPTCGKRVSLPNKALELKDRET 666
Db 333 LEGSPDKGLLPTKATKQKEDTSLGKLESPVXGDKLLKPTCGKRVSLPNKALELKDRET 387
QY 667 KVEEDFNLTKEGATKVTGQERDIGIIRAPQDQTNKMTSELGRKEDTKSDGEII 726
Db 388 -----PNKALELNEOTLRADILP-----SESKQDYESSWDSLS 425
QY 727 SVSDTONVECLPEATYQKEIKTTNGKIBES----- 756
Db 426 CETVSKDVLKPTKATKQKEDTSLGKLESPVXGDKLLKPTCGKRVSLPNKALELKDRET 485
QY 757 -----PEKSPHPEPATMONSVNPKGLEWKNQTLRA----- 788
Db 486 KAEPEKPSAFPAIEMQKSVNPKGLEWKNQTLRA----- 821
QY 789 -----DSTLSKILDALPSCERGERELKDNCEQITAKM 821
Db 546 ETVSQKDVCPKATKQKEDTSLGKLESPVXGDKLLKPTCGKRVSLPNKALELKDRET 605
QY 822 EQMKNFVQLKSELSAKIQLSENLQKAKWQELCSVRLPLNQBEERKRRNVDILKEKIR 881
Db 606 EQMKNFVQLKSELSAKIQLSENLQKAKWQELCSVRLPLNQBEERKRRNVDILKEKIR 665
QY 882 P-----EQLRKLEVKHOLEQTLRIODIELKSVTSNMQVSHTHESENDLPHECNMLKK 936
Db 666 BELGRIEEQHRELEVQKOLEQALRIQDIELKSVTSNMQVSHTHESENDLPHECNMLKK 725
QY 937 EIAMLKLEATLKHQVQENKVFEDIKILQENKAELOMTLKLKQKVTVKRASQYREOLK 996
Db 726 EIAMLKLEATLKHQVQENKVFEDIKILQENKAELOMTLKLKQKVTVKRASQYREOLK 785
QY 997 VLTAENTMLTSKLE 1011
Db 786 VLTAENTMLTSKLE 800

RESULT 15
AAU33357
ID AAU33357 standard; protein; 1095 AA.
XX
AC
XX
XZ
Dt 18-DEC-2001 (first entry)
XX
DE Human breast cancer protein B726P fusion protein #1.
XX
KW Human; ss; breast cancer protein; tumour; cancer; cytostatic;
KW gene therapy; immunogen.
XX
OS Homo sapiens.
XX
FN WO200179286-A2.
XX
PD 25-OCT-2001.
XX
PF 12-APR-2001; 2001WO-US012164.
XX
PR 17-APR-2000; 2000US-00551621.
PR 08-JUN-2000; 2000US-00590751.

22-JUN-2000; 2000US-00604287.
20-JUL-2000; 2000US-00620405.
(CORI-) CORIXA CORP.
Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
WPI; 2001-611721/70.
N-PSDB; AAS47421.

Breast Tumor Proteins and nucleic acids useful for the prevention,
diagnosis and treatment of breast cancer.
Claim 22; Page 292-295; 297pp; English.

The invention relates to isolated breast tumour proteins and nucleic
acids that encode them, including immunogenic fragments of the proteins.
Also included are expression vectors expressing the proteins, transformed
cells and antibodies raised against the proteins or an antigen presenting
cell expressing the protein. The proteins and nucleic acids may be used
in the prevention, diagnosis and treatment of diseases associated with
inappropriate breast tumour protein expression, i.e. breast tumours and
breast cancer e.g. by gene therapy. The nucleic acids and their
complements may also be used as DNA probes in diagnostic assays to detect
and quantitate the presence of similar nucleic acids in samples, and
therefore which patients may be in need of restorative therapy. The
proteins, nucleic acids and antibodies may be used in assays to identify
modulators (e.g. antagonists) of breast tumour protein expression and
regulate expression and activity. The antibodies may also be used as
diagnostic agents for detecting the presence of the proteins in samples
(e.g. by enzyme linked immunosorbent assay (ELISA)) and in other immuno-
purification diagnostic techniques. The present sequence is a breast
tumour protein encoded by a cDNA from a breast tumour cDNA library
isolated by subtractive hybridisation against a normal breast cDNA
library

Sequence 1095 AA;
Query Match 45.9%; Score 2372.5; DB 4; Length 1095;
Best Local Similarity 52.2%; Pred. No. 1.2e-151;
Matches 529; Conservative 73; Mismatches 151; Indels 261; Gaps 13;

QY 104 LLEHRLKLPKPNQNTDEGTSTGTGPDAAFLAERTPTAELSLLEKTPDEAARLVEGTSK 163
Db 24 LLENVISKTINPQVSKTE-----YKELQEFIDDNA-----TTNA 58
QY 164 IQCLGKATSGKFEQSTETPRKI-----LRPTKETSEKFSWPAKERSRK 207
Db 59 IDELKEC-----FLNQDTLNSVEFMQLIYDSSLCDLFMSPAKETSEKFTWAAGRPRK 114
QY 208 ITWEEKTSVKTECVAGVTPNKTVELEKGTSNMIACTKSTKASTNVDVSSVEPIFSL 267
Db 115 IAWKKETPVKTGCVARVTSNKTVELEKGRSKMIACTKSTKASTN----- 162
QY 268 PGTRTIENSQCTKVEEDFNLATKLIISKAAQNYCLPDATYQDKDKITINHKKIDQMPFSE 327
Db 163 -----DQRPPE 169
QY 328 SKREDEBEYSWDSGLSFESSAKTQVCIPESMYQKVMENRVEBELPEKPSAFKAVENOK 387
Db 170 SKQDEEYSCDSRSLFESSAKIQVCIPESYQKVMENRVEBELPEKPSAFKAVENON 229
QY 388 TVPNKAFELKNEOTLRAAQMPSPSKDDEENSWDSESPCETYSQKDVLPKATHQKEF 447
Db 230 SVNPKAFELKNEOTLRAAQMPSPSKDDEENSWDSESPCETYSQKDVLPKATHQKEI 289
QY 448 DTLGKLEESPVKQGLLPTKATKQKEDTSLGKLESPVXGDKLLKPTCGKRVSLPNKALELKDRET 507
Db 290 DKINGKLEESPVKQGLLPTKATKQKEDTSLGKLESPVXGDKLLKPTCGKRVSLPNKALELKDRET 349
QY 508 LPNKALELKDRETILKAEFPDNDGLLPTKATKQKEDTSLGKLESPVXGDKLLKPTCGKRVSLPNKALELKDRET 567
Db 508 LPNKALELKDRETILKAEFPDNDGLLPTKATKQKEDTSLGKLESPVXGDKLLKPTCGKRVSLPNKALELKDRET 567
```

Db 350 VPKALELKNBQTLRAD-----EILPSESKQ 375
Qy 568 KDEENSWDFESFLETLQNDVCLPKATHQKEFDTLGKLESPDKDGLLKPTCGMKISL 627
Db 376 KDYEESWDSSELCEVTSQKDVCLPKAXHQEIDKINGKLEGPVKDGLLKANCCKVSI 435
Qy 628 PNKALELKDRETFKAEDVSSVESTSLFGKPTTENSQSTKVEEDFNLTKEGATKTVTGO 687
Db 436 PTKALELMDMTFKAEPPK-PSAF-----EPAIEMQKSV-----PNKALELK 477
Qy 688 QERDIGIITERAQDQTNKMTPELGRKEDTSTSEIISVSDTONVECLPEATYQKEIK 747
Db 478 NEQTLRADEILP-----SESKQDYESSWDSSELCEVTSQKDVCLPKAXHQEID 528
Qy 748 TTNGKIEES-----PEKPSHFEPATEMNSV 773
Db 529 KINGLEESPDNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPKPSAFPAIEMQKSV 588
Qy 774 PNKGLEWQNKQTLRA-----788
Db 589 PNKALELKNBQTLRADQMPFSESOKQVSENSWDSSELRETYSQKDVCPKATHQKEMDK 648
Qy 789 -----DSTTLISKILDALPSCERGERELKDNCEQITAKMEQMKNFVLOKELSEAKEIK 842
Db 649 ISGKLEDSTLSKILDVHSCERARELODCEQRTGMEQMKKFCVLKXKLSAKEIK 708
Qy 843 SOLENQKAKWEQELCSVRLPLNQEEKERNVDILKEKIRP-----EEQLRKKLEVYHQLE 897
Db 709 SOLENQKVKWEQELCSVRLTLNQEEKERNADILNEKIRBELGRIEEQHREKLELVKQOLE 768
Qy 898 QTLRIQDIELKSVTSNLNOVSHTHESENDLPHENCMKKEIAMLKLEVATLKHQHVKEN 957
Db 769 QALRIQDIELKSVESNLNOVSHTHENENYLLHENCMKKEIAMLKLEIATLKHQVQKEN 828
Qy 958 KYFEDIKILQEKNAELOMTLXKOKTVTKRASQYREQLKVLTAENTMLTSKLKE 1011
Db 829 KYFEDIKILKERNAELOMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLKE 882

Search completed: February 19, 2005, 02:04:50
Job time : 117.137 secs

This Page Blank (uspto)

Db 181 CGVTAHYAVTCGPHIHEQIMEYIRKLSKNHNTNPEGTSAGTPDEAPLAERTPTDAE 240
Qy 241 SLVEKTPDEAPLVERTPDPAESLVEKTPDEASLVEGTSKIQCLEKATSGFEQSAE 300
Db 241 SLVEKTPDEAPLVERTPDPAESLVEKTPDEASLVEGTSKIQCLEKATSGFEQSAE 300
Qy 301 TPREITSPAKETSEKFTWPAKGRPRKIAWEKEDTPREIMSPAKETSEKFTWPAKGRPRK 360
Db 301 TPREITSPAKETSEKFTWPAKGRPRKIAWEKEDTPREIMSPAKETSEKFTWPAKGRPRK 360
Qy 361 IAWKKEPTPVKTCVARVTSNKTIVLEKGRSKMTACPTKSSSTKASANDORFSESQEE 420
Db 361 IAWKKEPTPVKTCVARVTSNKTIVLEKGRSKMTACPTKSSSTKASANDORFSESQEE 420
Qy 421 DEEYSCDSRSIFESAKIQVCPESIQYKWEINREVEEPPKPSAFKPAIEMQNSVPNK 480
Db 421 DEEYSCDSRSIFESAKIQVCPESIQYKWEINREVEEPPKPSAFKPAIEMQNSVPNK 480
Qy 481 AFEKNEQTLRADPMFPPEKQDYENSWSDESISLCTVSQKDVCLPKATHQKEIDKING 540
Db 481 AFEKNEQTLRADPMFPPEKQDYENSWSDESISLCTVSQKDVCLPKATHQKEIDKING 540
Qy 541 KLESPNKDGLLKATCGMKVSIPTKALELMDQTFKABPPGKPSAFEPATEMQKSVNKA 600
Db 541 KLESPNKDGLLKATCGMKVSIPTKALELMDQTFKABPPGKPSAFEPATEMQKSVNKA 600
Qy 601 LELKNEQTRADEIILPSESQKDYENSWSDESISLCTVSQKDVCLPKAAHQKEIDKINGK 660
Db 601 LELKNEQTRADEIILPSESQKDYENSWSDESISLCTVSQKDVCLPKAAHQKEIDKINGK 660
Qy 661 LEGSPVKDGLLKANGCMKVISIPTKALELMDQTFKABPPGKPSAFEPATEMQKSVNKA 720
Db 661 LEGSPVKDGLLKANGCMKVISIPTKALELMDQTFKABPPGKPSAFEPATEMQKSVNKA 720
Qy 721 ELKNEQTLRADEILPSESQKDYENSWSDESISLCTVSQKDVCLPKATHQKEIDKINGKL 780
Db 721 ELKNEQTLRADEILPSESQKDYENSWSDESISLCTVSQKDVCLPKATHQKEIDKINGKL 780
Qy 781 EBSPPNDGFLKAPCRMKVISIPTKALELMDQTFKABPPGKPSAFEPATEMQKSVNKALE 840
Db 781 EBSPPNDGFLKAPCRMKVISIPTKALELMDQTFKABPPGKPSAFEPATEMQKSVNKALE 840
Qy 841 LKNEQTLRADQMPPEKQDYENSWSDESISLCTVSQKDVCPKATHQKENDKISGKLE 900
Db 841 LKNEQTLRADQMPPEKQDYENSWSDESISLCTVSQKDVCPKATHQKENDKISGKLE 900
Qy 901 DSTLSKILDTVHSCERARELOKHCEQRTGKMEQMKKFCVLKXKLSAKIKSLEQLENO 960
Db 901 DSTLSKILDTVHSCERARELOKHCEQRTGKMEQMKKFCVLKXKLSAKIKSLEQLENO 960
Qy 961 KVKWEQELCSVRLTLNOEERKNADILNEKIREELGRIBEEQHRKELEVKQOLEQALRIQ 1020
Db 961 KVKWEQELCSVRLTLNOEERKNADILNEKIREELGRIBEEQHRKELEVKQOLEQALRIQ 1020
Qy 1021 DIEKSVESNLNQVSHTHENYLLHNCMLKEIAMLKLEIATLKHQYQENKIFYEDI 1080
Db 1021 DIEKSVESNLNQVSHTHENYLLHNCMLKEIAMLKLEIATLKHQYQENKIFYEDI 1080
Qy 1081 KILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLEKQDKKEILBAEIE 1140
Db 1081 KILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLEKQDKKEILBAEIE 1140
Qy 1141 SHPRLASAVODHDQIVTSRKSQBPAPHIAGDACLQKMNVDVSSITLYNNEVLHQLPSEA 1200
Db 1141 SHPRLASAVODHDQIVTSRKSQBPAPHIAGDACLQKMNVDVSSITLYNNEVLHQLPSEA 1200
Qy 1201 ORKSKSLKINLYAGDALRENTLVSEHAQORDQRETOCQKMEABEHYQNEQDNVNHKTEQQ 1260
Db 1201 ORKSKSLKINLYAGDALRENTLVSEHAQORDQRETOCQKMEABEHYQNEQDNVNHKTEQQ 1260
Qy 1261 ESLDQKLFQLOSKNMTLOQQVLVHAHKADNKSKITIDIHFLERKQOHLLKERNNEEFNY 1320
Db 1261 ESLDQKLFQLOSKNMTLOQQVLVHAHKADNKSKITIDIHFLERKQOHLLKERNNEEFNY 1320

Qy 1321 NNHLKNRIYOYEKEKAETENS 1341
Db 1321 NNHLKNRIYOYEKEKAETENS 1341
RESULT 2
Q9BX2 ID Q9BX2 PRELIMINARY; PRT; 1011 AA.
AC Q9BX2;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Breast cancer antigen NY-BR-1.1 (Fragment).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21174979; PubMed=11280766;
RA Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karbach J.,
RA Jaeger E., Knuth A., Old L.J., Chen Y.T.;
RT "Identification of a tissue-specific Putative Transcription Factor in
RT Breast Tissue by Serological Screening of a Breast Cancer Library.";
RL Cancer Res. 61:2055-2061(2001).
DR EMBL; AF269088; AAK27326.1; -;
DR HSSP; O75832; LUOH.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ANK; 3.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS00088; ANK REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW ANK repeat. 1011 1011
FT NON_TER 1011 1011
SQ SEQUENCE 1011 AA; 114250 MW; 16627D7B218DB438 CRC64;
Query Match 44.4%; Score 3075; DB 2; Length 1011;
Best Local Similarity 57.4%; Pred. No. 8.3e-104;
Matches 664; Conservative 76; Mismatches 145; Indels 272; Gaps 12;
Qy 98 VVAKLLSHGAVIEVHNKASLTPLLLSITKRSQEVFELLKNANANVKNYKCTALMLAV 157
Db 1 MVATLLSYGAVIEVQNKASLTPLLLAIQKRSQVFEVLLTNANANAFNSKCTALMLAI 60
Qy 158 CHGSEIYGMLLQONVDVFAADICGVTAHYAVTCGPHIHEQIMEYIRKLSKNHNTN 217
Db 61 CEGSEIYGMLLQONVDVFAEDIHGITAERYAAARGVNYHQQLLEHIRKLPKPNQNTN 120
Qy 218 ECTSAGTPDEAPLAERTPTDAESLVEKTPDEAPLVERTPDPAESLVEKTPDEASLVE 277
Db 121 EGTSTGTDPDEAPLA-----BRTPTDAESLVEKTPDEAPLVE 158
Qy 278 GTSKDIQCLEKATSGFEQSAEETPREITSPAKETSEKFTWPAKGRPRKIAWEKEDTPR 337
Db 159 GTSAKIQCLGKATSGFEQSTETPRKILRPYKETSEKSWP----- 200
Qy 338 EIMSPAKETSEKFTWAAKGRPRKIAWEKETPVKTCVARVTSNKTIVLEKGRSKMTAC 397
Db 201 -----AKERSRKITWEEKETSVKTECVAGVTPNKTEVLEKGTSSNMIACP 244
Qy 398 TRESSTKASAN----- 408
Db 245 TKETSTKASTNVDVSVSEPIFSLFGTRTTIENSQCTKVEEDFNLATKIISKSAQNYTCLP 304
Qy 409 -----DQRFPPSEKOEDEEYSCDSRSIFESAKIQVCPESIQYKWE 452
Db 305 DATYQDKITNKIEDQMFPSSEKEEEDYSCDSRSIFESAKIQVCPESIQYKWE 364
Qy 453 INREVEEPPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADPMFPPEKQDYENSWS 512
Db 365 INREVEEPPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADPMFPPEKQDYENSWS 424

```
Qy 513 ESLCETVSQDVCLPKATHQKEIDKINGKLBESPNVOGLLKATCGMKVSIPTKALELMD 572
Db 425 ESPCETVSQDVCLPKATHQKEIDKINGKLBESPNVOGLLKATCGMKVSIPTKALELMD 484
Qy 573 QTFKAEPKPSAFEP-----ATEMQK-----SYPN 598
Db 485 ETFKAEPKPSAFEP-----ATEMQK-----SYPN 544
Qy 599 KALELKNQWRADDEILPESKQDYENSWDTSCLCETVSQDVCLPKAQAQKEIDKIN 658
Db 545 KALELKNQWRADDEILPESKQDYENSWDTSCLCETVSQDVCLPKAQAQKEIDKIN 604
Qy 659 KLESPKQDVCLPKATHQKEIDKINGKLBESPNVOGLLKATCGMKVSIPTKALELMD 713
Db 605 KLESPKQDVCLPKATHQKEIDKINGKLBESPNVOGLLKATCGMKVSIPTKALELMD 664
Qy 714 SV-----PNKALELKNQWRADDEILP-----SESKQDYENSWDSE 751
Db 665 STKVEEDPNLTKEGATKVTGQQRDGIITRAPQDQTNKMTSELGRKEDTKSTDSSE 724
Qy 752 SLCETVSQDVCLPKATHQKEIDKINGKLBESPNVOGLLKATCGMKVSIPTKALELMD 811
Db 725 IISVSDIQNYECLPEATYQKEIKTTNGKIES----- 756
Qy 812 TFKAEPKPSAFEPAPTEMQSVNPKALELKNQWRADDEILPESKQDYENSWDSE 871
Db 757 -----PEKPSHFEPATEMQSVNPKALELKNQWRADDEILPESKQDYENSWDSE 788
Qy 872 LRETVSQDVCLPKATHQKEIDKINGKLBESPNVOGLLKATCGMKVSIPTKALELMD 931
Db 789 -----DSTLTKSLDALPSCRGRLKQDCEQITA 819
Qy 932 KMEQKKKFCVLKLLSEAKIKSLENQKQVKEQELCSVRLTLNQBEKERNADIINEK 991
Db 820 KMEQKKKFCVLKLLSEAKIKSLENQKQVKEQELCSVRLTLNQBEKERNADIINEK 879
Qy 992 IREELGRIEBQHRKEVQKQLEALRIQDIELKSVEENLNQVSHENNYLLHENCML 1051
Db 880 IRP-----EQURKKLEVKHOLEQTLRIQDIELKSVEENLNQVSHENNYLLHENCML 934
Qy 1052 KKEAMUKLEATLKHQVQENKVFEDIKILKEKNAELOMTLKEESLTKRASQYSGQ 1111
Db 935 KKEAMUKLEATLKHQVQENKVFEDIKILKEKNAELOMTLKEESLTKRASQYSGQ 994
Qy 1112 LKVLIAENTMLTSKLKE 1128
Db 995 LKVLIAENTMLTSKLKE 1011

RESULT 3
Q9H1Q1 PRELIMINARY, PRT, 1710 AA.
ID AC Q9H1Q1
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BA145E8.1 (KIAA1074).
GN Name=BA145E8.1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Bird C.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL162272; CAC19649.1; -.
DR HSSP; P20749; 1K1B.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ANK; 4.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS00086; ANK_REPEAT; 4.
```

```
DR PROSITE; PS00297; ANK_REPEAT; 1.
KW ANK repeat.
SQ SEQUENCE 1710 AA; 196409 MW; 01CBF9BADB894872 CRC64;

Query Match 25.1%; Score 1740; DB 2; Length 1710;
Best Local Similarity 33.8%; Pred. No. 3.2e-55;
Matches 479; Conservative 204; Mismatches 376; Indels 358; Gaps 34;

Qy 10 LNIQDAQKRTALHWACVNGHEEVVTLVDKRCQDLVDLGEHRTPLMKALQCHQACANTL 69
Db 73 LNDKQKRTALHLACANGHEEVVTLVDKRCQDLVDLGEHRTPLMKALQCHQACANTL 132
Qy 70 IDSGADINLVYGNMAYVSEILSVAKLSHGAVIEVHNKASITPLLSITKRSSE 129
Db 133 LEHGADPNLADVHGNTALHYAVYVNEISVATKLLLYDANIEAKNKDLDLTPLLILAVSGKKQ 192
Qy 130 QIVFELLIKANANAVNKYKCTALMLAVCHSGSSEIVGMLLQNVNDVPAADICGVTAHYA 189
Db 193 QMVFEFLIKKANVANVADKLESS----- 214
Qy 190 VTCGFPHIHEQIMYIRKLSKNHONTNPEGTSAAGTDEAALPLAERTPTDAESLVEKTPDE 249
Db 215 -----HQLISYKEERIPKHSQNSNSVDESEDSLSRLSGK-PGV-----DD 256
Qy 250 AAPLVERTPTDAESLVEKTPDEAASLVGTSKIQCLEKATSGKGFQSAEETPR--RITS 307
Db 257 SWP-----TSDD-----EDLNFDTKQVVPKPSLAKMLMTASQOSRKNLEATY 296
Qy 308 PAKETSEKFTWPAKGRPRKIAWEKKEPTPRIMSPAKETSEKFTWAAKGRPRKIAWEKKE 367
Db 297 GTVRTGN-----RTLFEDRDSQDEVVVESELT----- 325
Qy 368 TPVATGCVARVTSNKTQVLEKGRSMIACP--TKESSTKASANDQRFPSSEKQEBDEYS 425
Db 326 TSIKQVCFESHYTSYDPLLPKPSHSLANPLGMLKEETK-----PGIAKENGID-- 376
Qy 426 CDSRSLFESSAKIQVCIPESYQKWEINREVEBPVKPSAFKPAIEHQNSVFNKAFELK 485
Db 377 -----IESAP-----LEQTN--DNLTYVDE 395
Qy 486 NEQTLRADPMPFPSPKQDYENSWDSESLCETVSQDV--CLPKATHQKEIDKINGKLEE 544
Db 396 VHKNNRSDMSALGLGQEEEDIESPWSSESISENFPQKYVDPLAGAADGKKNIGNEQAE- 454
Qy 545 SPNKDGLLKATCGMKVSIPTKALELKD-----MOTFKABPPKPSAFEPATEMQSVNPK 599
Db 455 -----DVFIYPC--MSGSENFMAKLEDTRNVMVAHMESEBYVHLKPTTEMKDSVFNK 509
Qy 600 ALELKNQWRAD-----EILPESKQDYENSWDTSCLCETVSQDVCLPKAQAQKEID 655
Db 510 AGMKDVQTSKAAEHDLVASEEEREGSEN-----NQPVVEERKKNHNM 558
Qy 656 KINGKL-----EGSPVKDGLLKANGCMKVSIPTKALELMDMOTFKABPEKPSAFEP 707
Db 559 EVSANIHDGATDDAEDDDDDGLIQXR-----KSGE--TDHQQFPRKENKEYASSGP 608
Qy 708 AIEQKSVNPKALELKNQWRADDEILPESKQDYENSWDTSCLCETVSQDVCLPKAQAQKEID 767
Db 609 ALQNK-----EVKSTEKEKTSKES----- 628
Qy 768 THQKEIDKINGKLBESPNVOGLLKATCGMKVSIPTKALELMDMOTFKABPEP---KPSAF 824
Db 629 -----VNSPVFGKASLLTGGLLQVDDSSLSLSEIDEGRPT-- 664
Qy 825 EPAIEMQSVNPKALELKNQWRADQMPFSPSEKQKVEENSWDSESLRETVSQDVCPV 884
Db 665 -----KTSNEKNKVNQIQSMDD-----VDDLTSQSETASE-----DCLSLP 701
Qy 885 KATHQKEM---DKISGLEKEDSTLSKILDTVHSCERARELOKDHCEQRTKMEOMKKKFC 941
Db 702 HSSYNNFMILLIEQLGMECKDSVLLKQDAALSCEKRLLEKKNHCELLTIVLKMEKDVN 761
Qy 942 VLKKKLSEAKIKSLENQKQVKEQELCSVRLTLNQBEKERNADIINEKIREELGRIEE 1001
```

```

Db 762 VLQRELSETEIKSLEHVKQVEMRELCSEFSLNQBEEKRNADTLYEKIRBQLRRKEE 821
Qy 1002 QHRELEVKQOLEALRIQDIELKSVSNLNOV-----1034
Db 822 QYRKEVEVKQOLELSQLELMEKLVKSNLNOVQVQERNDAQRLSRQNAARMLODGLTN 881
Qy 1035 -----SHTHENENYLLHNCMLKKEIAMLKLEIATLKHQYOEKENK 1075
Db 882 HLSKQKETEMAKKMSNSHSEHEEDKSHKSNLQBEIATMLRLIEDITKNQNOKEKK 941
Qy 1076 YFEDIKILKEKNAELQMLTKLKEBSLYKRAOYSGQLKVLTAENTMLTSKLK-EKQDKEI 1134
Db 942 CFEDLKIKVEKNEDLQKTIKQNEETLQTIISQYNGRLSVLTAENAMLSKLENEKQKER 1001
Qy 1135 LEAEETHHPRILASAVODHQLVTSRKQSEPAFIAGDAC--LQKMNVDVSYSTIYNNEV 1192
Db 1002 LEAEVESTHSLAAIHDRDQSETSKRELELAFQFARDECRLQDKMNFVSNLKDNEI 1061
Qy 1193 LHQPLSEARQSKSLKINLVAGDALRENTLVSEHAQRDQRETQCKMEAEHMYQNEQDN 1252
Db 1062 LSQQLFKTESKLSLEIEFHTRDALREKTLGLERVOXDLSQTOCKMEAEQYQNEQVK 1121
Qy 1253 VNKHQEBSLDOKLFOLOSKNMMLQOOLVHAHKADNKSITIDI----HFLERKMOHH 1308
Db 1122 VNKYIGQSEVEERLSQLESENMLRQQLDADHAKADNKEKTVINIQDFHAIQVKLOAE 1181
Qy 1309 -----LLKEKNEEIEFNYNHKLRIYQVEKEKAETE 1339
Db 1182 SEKQSLLEERKNELISECNHLKEXQYQYENKAEKE 1218

```

RESULT 4

```

QUP88 PRELIMINARY; PRT; 1715 AA.
AC QUP88;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2004 (TREMBLrel. 26, Last sequence update)
DE 01-MAY-2004 (TREMBLrel. 26, Last annotation update)
DE KIAA1074 protein (Fragment).
GN Name=KIAA1074;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99397452; PubMed=10470851;
RA Kikuno R., Nagase T., Ishikawa K., Hirose M., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.
RT "Prediction of the coding sequences of unidentified human genes. XIV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res 6:197-205(1999).
DR ENBL; AB028997; BA83026.2; -.
DR HSP; P20749; IkiB.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 4.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SMC0248; ANK; 4.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW ANK repeat.
FT NON TER
SQ SEQUENCE 1715 AA; 196962 MW; A82858F5F58E3203 CRC64;

Query Match 25.1%; Score 1736.5; DB 2; Length 1715;
Best Local Similarity 33.9%; Pred. No. 4.3e-55;
Matches 481; Conservative 200; Mismatches 377; Indels 359; Gaps 35;

Qy 10 LNIQDAQKRTALHWACVNGHEVTVFLVDRKCOLDVLGDGHRPLMKALQCHQACANIL 69

```

```

Db 79 LNDRCQNRRTALHLACANGHPEVVTLLVDRKCOLQVNDNENRTALMKAVQOEKCATIL 138
Qy 70 IDSGADINLVGYGNMHLVAVYSEITLTVVAKLSHGAVIEVHNKASLTPLLLSITKRSE 129
Db 139 LEHAGDPNLDVHGTALHYAVYNEIDISVATKLLLYDANIEAKNKDDLTPLLLAVSGKKQ 198
Qy 130 QIVFELLIKNANANVNYKTKTALMLAVCHGSEIIVGMLLQQNVVFAADICGVTAHYA 189
Db 199 QMVEFLIKKANVNAVDKLESS-----220
Qy 190 VTCGFHHIHEIMYIRKLSKNQNTNPEGTSAGTPEAAPLAERTPTDAESLVEKTPDE 249
Db 221 -----HOLISEYKEERIPKHSQNSNSVDESDLSRUSGK-PGV-----DD 262
Qy 250 AAPLVERTPTAESLVEKTPDEAASLVEGTSKIQCLEKATSKGFKQSABETPR--EITS 307
Db 263 SWP-----TSDD-----EDLNFDTKNVPKSLAKMTASQOSRKNLEATY 302
Qy 308 PAKETSEKFTWPAKGRPKIAWEKEDTPREIMSPAKETSEKFTWAAKGRPKIAWEKKE 367
Db 303 GTVRTGN-----RTLFEDRDSQDEVVVEISLPT-----331
Qy 368 TPVKTGCVARTSNKTVLEKGRSKMTACP--TKESSTKASANDQRPFSKOBEDREYS 425
Db 332 TSIKVQCFSHTYQSPDLLPKPSHKSANPLMKBEPTK-----PGIAKKENGIDI- 382
Qy 426 CDSRSLFESSAKIQVCIPESIQYKVMIEINREVBEPPKPSAFKPAIEMQNSVNPNAFELK 485
Db 383 -----IESAP-----LEQTN-DNLTVYDE 401
Qy 486 NEQTLRADPMPFPSPESKQDYENSWDSSELCTVSOQDV-CLPKATHQKEIDKINGKLEE 544
Db 402 VHKNRSDMMSALGLQGEEDIESPDSEISSENPQKYVDPLAGAADCKEKNIGNEQAE- 460
Qy 545 SPNKDGLLKATCGMKVSIPTKALELKD-----MOTFKAEPPKPSAFEPATEMOKSVPNK 599
Db 461 ----DVFIYFSC-MSGSRNFQAKLETRNVGMPVAHVESPERVHLKPTIEMKDSVPNK 515
Qy 600 ALELKNQETRAD----EILPSESKQDYENSWDTSSELCTVSOQDVCLPKAAHQKEID 655
Db 516 AGMKQDVQTSKAAEHDLVASEEBOERESN-----NQPOVEERKKHNNEM 564
Qy 656 KINGKL-----EGSPVKDGLLKANGCMKVSIPTKALELMDMTQFKAEPPKPSAPEP 707
Db 565 EVSANIHGDGATDDAEDDDDDGLIQKR-----KSGE-TDHOQF-----P 602
Qy 708 ALEMOKSVPNKALELKNQETLRADIELPSPESKQDYESSWDSSELCTVSOQDVCLPKA 767
Db 603 RKENKEYASGALQNK-----EVKSTEKEKTSKES-----633
Qy 768 THQKEIDKINGKLEESPDNGFLKAPCKMKVSIPTKALELMDMTQFKAEPPPE--KPSAF 824
Db 634 -----VNSPVFGKASLLTGLGLQVDDSDSLSEIDEDEGRPT--669
Qy 825 EPAIEMOKSVPNKALELKNQETLRADQMPFSPESKQDYENSWDSSELCTVSOQDVCLPK 884
Db 670 -----KKTSEKKNKVNQIQSMDD-----VDDLTSQSETASE-----DCELP 706
Qy 885 KATHQKEM---DKISGKLEDSTLSKILDTVHSCERARELOKHCEQRTGKMEQWKYKFC 941
Db 707 HSSYKNFMILLIQLGMECKDSVSLIKIQAALSCERLLEKKNHCELLTVTKIKMEDKYN 766
Qy 942 VLKCKLSEAKETIKSQLENQKQVWEOELCSVRLTINQOEKRENADILNEKIEELGRIBE 1001
Db 767 VLQRELSETEIKSLEHVKQVEMRELCSEFSLNQBEEKRNADTLYEKIRBQLRRKEE 826
Qy 1002 QHRELEVKQOLEALRIQDIELKSVSNLNOV-----1034
Db 822 QYRKEVEVKQOLELSQLELMEKLVKSNLNOVQVQERNDAQRLSRQNAARMLODGLTN 886
Qy 1035 -----SHTHENENYLLHNCMLKKEIAMLKLEIATLKHQYOEKENK 1075
Db 887 HLSKQKETEMAKKMSNSHSEHEEDKSHKSNLQBEIATMLRLIEDITKNQNOKEKK 946

```

```

QY 1076 YFEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSLKJL-EKQDKFI 1134
DB 947 CFEDLKIVKQKNEEDLQTKIKONEETLTQISQYNGRLSVLTAENAMLSKLENEKQSKER 1006
QY 1135 LEAETESHPRILASAVQDHQIVTSRKSQEPAFHAGDAC--LQKMNVDVSTIYNNEV 1192
DB 1007 LEAEVESHRSRLAALHQRDSQETSRLLELAFQARDECSRLQDKQNFVSNLKNKNEI 1066
QY 1193 LHQPLSEAKQSKSLKINLYAGDALRENTLVSEHAQRDQRETQCCQKAEAHMYQNEQDN 1252
DB 1067 LSQQLFKTESKLSLEIEFHHTRDALREKTLGLERVQKLSQTCQCKEMEQKYNEQVY 1126
QY 1253 VNKHTEQESLSQKLPOLQSKNMWLOQLVHAHKADNKSITIDI----HFLERKMHQ 1308
DB 1127 VNKYIGKQSVBERLSQSENMFLRQQLDDAHNKADNKEKTVINIQDQFHAIVQKQAE 1186
QY 1309 -----LLKEKNEEFYNNHLKNRIYQYEKEKAETE 1339
DB 1187 SEKQSLLEERNKELISECNHLKEROYQYENKAE 1223

RESULT 5
Q9NSI9
ID Q9NSI9 PRELIMINARY; PRT; 424 AA.
AC Q9NSI9
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DE PRE04 protein (Fragment).
GN Name=PRE04;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.S.,
RA Toyoda A., Ishii K., Totoki Y., Choi D.K., Soeda E., Ohki M.,
RA Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudon J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,
RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Leinhardt R., Raspo M.B.;
RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL163203; CAB90394.1; -.
FT NON TER 1
FT NON TER 424
SQ SEQUENCE 424 AA; 50038 MW; C6E704795C534625 CRC64;

Query Match 24.8%; Score 1718; DB 2; Length 424;
Best Local Similarity 79.7%; Pred. No. 4.8e-55;
Matches 350; Conservative 30; Mismatches 43; Indels 16; Gaps 4;

QY 901 DSTLSKILDTVHSCERARELOKDHCEQRTGQEQKMKKFCVLKKLSEAKIKSOLNQ 960
DB 2 DSTLSKILDTVHSCERARELOKDHCEQRTGQEQKMKKFCVLKKLSEAKIKSOLNQ 61
QY 961 KVKVQELCSVRLTNQBEKERNADILNEKREELGRIEEHRKEVLEKQLEALRIQ 1020
DB 62 KAKVQELCS-----KKRDVDILKEKIRP-----EEQRLKKEVLEKQLEALRIQ 107
QY 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKQYQENKRYFEDI 1080
DB 108 DIELKSVTSLNQVSHTHESNDLFHENCMLKKEIAMLKLEIATLKQYQENKRYFEDI 167
QY 1081 KILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSLKJL-EKQDKFI 1140

```

```

DB 168 KILQEKNAELQMTLKLKQKTLTKRASQYREQQLKVLTAENTMLTSLKJL-EKQDKFI 227
QY 1141 SHHPRILASAVQDHQIVTSRKSQEPAFHAGDACLQKRNVDVSTIYNNEVHLHQLPSEA 1200
DB 228 SHHPRILASALQDHQDSVTSRKNQELAFHAGDAHLQGINVDVSNITIIYNNEVHLHQLPSEA 287
QY 1201 QRKSKSLKINLYAGDALRENTLVSEHAQRDQRETQCCQKAEAHMYQNEQDNVKNKTEQ 1260
DB 288 QRKSKSPKINLYAGDDLRENALVSEHAQRDQRETQCCQKAEAHMYQNEQDNVKNKTEQ 347
QY 1261 BSLDQKLPOLQSKNMWLOQLVHAHKADNKSITIDIHFLERKMHQHLKEKNEEFYNN 1320
DB 348 BSLDQKLPOLQSKNMWLOQLVHAHKADNKSITIDIHFLERKMHQHLKEKNEEFYNN 405
QY 1321 NNHLKNRIYQYEKEKAETE 1339
DB 406 GNHLKEDICQYEKEKAETE 424

RESULT 6
Q6ZR14
ID Q6ZR14 PRELIMINARY; PRT; 1080 AA.
AC Q6ZR14
DT 05-JUL-2004 (TremBLrel. 27, Created)
DT 05-JUL-2004 (TremBLrel. 27, Last sequence update)
DE Hypothetical protein FLJ46736.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kaneshiro K., Ishibashi T., Chiba Y., Fujimori K., Hirooka S.,
RA Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A.,
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.;
RA Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK128577; BAC87508.1; -.
FT NON TER 1080
SQ SEQUENCE 1080 AA; 126741 MW; 47BBA9333AE595C CRC64;

Query Match 18.0%; Score 1246; DB 2; Length 1080;
Best Local Similarity 41.2%; Pred. No. 1.6e-37;
Matches 307; Conservative 122; Mismatches 200; Indels 116; Gaps 16;

QY 678 KVSIPTKALELMDMOTFKAE-----PPEKPSAPEPAIMQKSVNKALELKN---EQLTRA 730
DB 64 KVSLEELQ--DMQRFKNEIGMLKVEFQALEKEKQVQKEVEERKRRNNEVMSANI 121
QY 731 DEILPSESQKQDYESSWDSLSCTVTSQKDVCLPK-----ATHQEKIDKNGK 779
DB 122 HDGATDDAEDDDDDGLQKRGSGTDDHQ---PFRKENKEYASSGPAALQMKVEKSTKE 178
QY 780 LEESPDNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPE---KPSAPEPAIMQKSVN 836
DB 179 KRTSKES---VNSPVFGKASLLTGGLQVDDSSLSLSEIDEDGRPT-----KKTNS 226
QY 837 KALELKNQTLRADQMPFPSESKQKVEENSWSLSRETSTVSQKDVCPKATHQKEM---D 893
DB 227 EKNYKNGIQSMDD-----VDLITQSSETASE-----DCLPHSSYKFMILLIE 270
QY 894 KISGKLESTLSKILDTVHSCERARELOKDHCEQRTGQEQKMKKFCVLKKLSEAKIK 953
DB 271 QLGMECKDSVLLKIQDAALSCERLLELKNKNCCELLTVKIKKMDKVNVLQELSETKEI 330
QY 954 KSOLENQKVKVEQLCSVRLTNQBEKERNADILNEKREELGRIEEHRKEVLEKQLEALRIQ 1013
DB 331 KSOLEHQKVEERELCSLRFSLNQBEKERNADTLTYEKIREQLRRKESQYRKEVEVKOOL 390

```

QY	1014	EQALRIQDIELKSVESNINQV	-----	1034			
DB	391	ELSLQTLMELETKVSNLNVQVQERNDAQRLSREQNARMLODGLITNHLKSQKEIEMAQ	450				
QY	1035	-----SHTHENENYLLHENCMLKETAMLEKLEITATLKHOYQKKNYFEDIKILKEKN	1087				
DB	451	KOMNSHSHSEEEKDLGSHKSMLOQBEIAMELRLEIDITKNQOEKEKCFEDLKIIVKKN	510				
QY	1088	AELOMTLKLKESLTKRASQYGGQLKVLIAENTMLTSKL-EKQDEILEABIESHPRL	1146				
DB	511	EDLOKTIKQNEFTLQITISQNGRLSVLTAENAMLSKLENEKQSKERLEAEVESHSL	570				
QY	1147	ASAVQDHQIVTSRKSQEPAFIAGDAC--LQKMNVDVSVSTIYNNEVLHQPLSAQRKS	1204				
DB	571	AAAIHDDQSETSKRELEAFQARDECRLQDKMNFVSNLKNNEILSQOLFETESKL	630				
QY	1205	KSLKINLNYAGDALRENTLVSEHAQDQRETQOCQKEAEHMYQNEQDNVKNKHTQOQESLD	1264				
DB	631	NSLEIEFHTRDALREKTLGLERVQKDLSTQOCQKEMEQTQONQVKNKYIGHQESVE	690				
QY	1265	OKLFQLOSKNMMLQOQLVHAHKADNKSITIDI-----HFLERKMQHH-----LLKEKN	1314				
DB	691	ERLSQLOSENMLRQQLDDAHKADNKEKTVINIQDFHAIYQKLQAESEKQSLLEERN	750				
QY	1315	BEIFNYYNNHLKNRIYOYEKEKAETE	1339				
DB	751	KELISECNHLKERQYOYENKAEARE	775				
RESULT 7							
QY	Q9H0H6	PRELIMINARY;	PRT;	823 AA.			
AC	Q9H0H6						
DT	01-MAR-2001	(T-EMBLrel. 16, Created)					
DT	01-MAR-2001	(T-EMBLrel. 16, Last sequence update)					
DT	01-MAR-2004	(T-EMBLrel. 26, Last annotation update)					
DE	Hypothetical protein DKFZp434A171.						
GN	Name=DKFZp434A171;						
OS	Homo sapiens (Human).						
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;						
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
OX	NCBI_TaxID=9606;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	TISSUE=Testis;						
RG	The German cDNA Consortium;						
RA	Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,						
RA	Mewes H.W., Weil B., Amid C., Oeanger A., Fobo G., Han M., Wiemann S.;						
RL	Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.						
DR	EMBL;	AL136793;	CAB66727.1;	-			
DR	HSSP;	Q60778;	10Y3.				
DR	InterPro;	IPR002110;	ANK.				
DR	Pfam;	PF00023;	ANK;	6.			
DR	PRINTS;	PR01415;	ANKYRIN.				
DR	SMART;	SM00248;	ANK;	6.			
DR	PROSITE;	PS50088;	ANK_REPEAT;	4.			
DR	PROSITE;	PS50297;	ANK_REPEAT_REGION;	1.			
KW	ANK repeat; Hypothetical protein.						
SQ	SEQUENCE	823 AA;	93968 MW;	9D2DDD5F52C379B2 CRC64;			
Query March 16.1%; Score 1112; DB 2; Length 823;							
Best Local Similarity 24.6%; Pred. No. 8.4e-33;							
Matches 329; Conservative 165; Mismatches 248; Indels 596; Gaps 21;							
QY	5	KKTINLNIQDAKRTALHWACVNGHEEVTVFLVDRKQOLDVLDEGHTPLMKALQCHOEA	64				
DB	55	RRSGDLALDQKRTALHLACASGVQVTLVNRKQCIDVCDKXENRTPLIAVHCQEEA	114				
QY	65	CANILIDSGADINLVGVNMLHYAVSYELSVVAKLISHGAVIEVHNKASLPLLSI	124				
DB	115	CAVILLERGANPNFKDIYGNLTHYAVSYESTSLAEKLLSHGAIHALDKDNTPLLFAI	174				

```
Db 666 ETQIFNALDKFCKTRSLRKSJLALETVDLSQTOQOQEMKEMYNABAKYNNSTG 725
Qy 1259 QOESLDQFQOLQSKNMILQOOLVIAHKKADNKSKIT-IDIHFLERKMOHLLKKEKNEI 1317
Db 726 KWCVEERICHLORENANVLQOQDDVHQEDHKEIVTNLQRFIESEKDLVLEKSKL 785
Qy 1318 FNNHNLKNRIYQYEK 1335
Db 786 MNECDHLKESLFQYEREK 803

RESULT 8
Q692S2
ID Q692S2 PRELIMINARY; PRT; 1043 AA.
AC Q692S2;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE MKIAA1074 protein (Fragment).
GN Name=MKIAA1074;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Seino S., Nishimura M., Kaisho T., Hoshino K., Kitamura H.,
RA Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene:
RT IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous
RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
RT Randomly Sampled from Size-Fractionated Libraries.";
RL DNA Res. 11:205-218(2004).
DR EMBL; AK173096; BAD32374.1; -.
DR InterPro; IPR010989; t-snare.
FR InterPro; IPR009054; Topoisomerase_insert.
FT NON TER 1
FT NON TER 1043
SQ SEQUENCE 1043 AA; 121354 MW; 5F37B1F9AA37B212 CRC64;

Query Match 13.9%; Score 962.5; DB 2; Length 1043;
Beat Local Similarity 35.2%; Pred. No. 2.8e-27;
Matches 252; Conservative 115; Mismatches 191; Indels 157; Gaps 17;

Qy 696 APPPEKPSAFEPKSAFPAIEMQKSVNPKALELNEQTLR-----ADSLPSESKQDYESS 747
Db 1 AEP--SGNDYGAADGGADVPKQSGDTENQQPSRGSGRGGPALLMKKAKMENEK-- 56

Qy 748 WDSLSLCTVSKQVCLPKATHQKIDKINGKLESPNDGFLKAPCRMKVSIPTKALEL 807
Db 57 WVSREPARTAMSGTGLP-----TGWPQMDGSCW----- 87

Qy 808 MDMTFKAEPPKPSAFEPKSAFPAIEMQKSVNPKALELNEQTLRADQM-----FPSESQKKVE 863
Db 88 SDTQSEARPTKTS-----SKNK-----DSGQTAAVDNLDLDFTESSETASEDHE 133

Qy 864 ENSWDSLSLRETQSKQVCPKATHQKMDKISGLKLESTLSKILDTVHSCERARELQK 923
Db 134 LQGPSESI-----LCA-----IEHURLCKDTASLTKLDVAVSYKRIELKR 177

Qy 924 DHCEQRTGMQMKKFCVKKLSEAKEIKSQLENQKVKQEQELCSVRLTLNQBEKKR 983
Db 178 SHCELLTGKLEKEMNKYGLQKEMSETEVKSLRHEKVGMEQELCLRFLFKQBEKKR 237

Qy 984 NADILNEKIRELGRIBEOHRKELEVKKOQLEALRIQDIELKSVSNLNVQ----- 1034
Db 238 SADQLSEKTMELRRKGECQSEVAREQLEASRLTLEMELTKVKSHLNVQLNEERETOR 297

Qy 1035 -----SH-----THENYLLHENCMLKKEIAM 1057
Db 298 QLSREQNARMLODGLASHLCKQKQEIEMTQKMTSEVSVSEKEDLLHKNQRLQDEAV 357
```

RESULT 9

```
Q9HCD1
ID Q9HCD1 PRELIMINARY; PRT; 718 AA.
AC Q9HCD1;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE KIAA1641 protein (Fragment).
GN Name=KIAA1641;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVIII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro.";
RL DNA Res. 7:273-281(2000).
DR EMBL; AB046861; BAB3467.1; -.
FT NON TER 1
FT NON TER 1
SQ SEQUENCE 718 AA; 82972 MW; 2AD05ADE70E3C97F CRC64;

Query Match 13.1%; Score 906; DB 2; Length 718;
Beat Local Similarity 35.8%; Pred. No. 2.1e-25;
Matches 241; Conservative 131; Mismatches 237; Indels 64; Gaps 16;

Qy 700 EKPSAFEPKSAFPAIEMQKSVNPKALELNEQTLRADEILPSESQKDYESSWDSLSLCTVQ 759
Db 13 QKQPALKATSKDSVSNIPTEIKDQOQ-----SGTVSSQK---QPAWKATSV-----K 58

Qy 760 KDVLCPKATHQKIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDMTQFKAEPP 819
Db 59 KDSVSNIAATEIKD-QQIRGTV--SPQKSAQKVIKKKVSLLNIAITRITGKWSGTEYP 115

Qy 820 KPSAFEPKSAFPAIEMQKSVNPKALELNEQTLRADQ--MFPSESQKDYESSWDSLSRETYS 877
Db 116 NLPLTKATYENKNSVINTATKMKDQVTSTPQDLEMAEGEQKLEEVNNOQVKNQTH 175

Qy 878 OK---DVCVPKATHQK-----MDKISGLKLESTLSKILDTVHSCERAR 919
Db 176 SRDLDLDLIIQSQTVESDGLCCNCKNVILLIDHEMKCKDCVHLKIKNTFCLWKRLI 235

Qy 920 ELQKDHCBQRTGRMBQMKKFCVLKKLSEAKEIKSQLENQKVKQEQELCSVRLTLNQBE 979
Db 236 KLNHCQELQVIRKLNKASVLQKRISEKEIKSQLEKELCSLRFALQDQSK 295

Qy 980 EKRNRADILNEKIRELGRIBEOHRKELEVKKOQLEALRIQDIELKSVSNLNVQSHVTE 1039
```


Db 296 KKRNVVELHOKVREKLITEEQRIEADVTKPIKPAKLSABVELTKGNNNSQVSETDE 355
 QY 1040 NENYLLHNCMLKKEIAMLKLEIATLKHOYQKQENKYPEDIKILKKAELQMTLKKEE 1099
 Db 356 KED-LLHENRLMQDEIARLEKDTIKQNLK--KYLKDFEIVKREHEDLQKALKRNGE 412
 QY 1100 SITKRASQVSGOLKVLIAENTMLTSL-KEKODKILAEATESHPRLASAVQDHDQIVT 1158
 Db 413 TLIAKTIACVSGQALATDENTTLRSKLEKQORLETEMOSYICRLNAARCDHQSHS 472
 QY 1159 SRKSQEPAPHIAQDAC--LQRKXNVDSVTIYNNEVLHQLPSEAQRKSKSLKININAYAD 1216
 Db 473 SKEDQELAFQGVKCRHLQENLNSHL-----ILSLQSKAASKSRVLKTELHYTGE 525
 QY 1217 ALRENTLVSEHAQRORETOCQKAEAHMYQNEQNVNKHTEQOESLQKLPOLQSKNMW 1276
 Db 526 ALKERALVFEHVQSELKQSQMKDIKMYKSYNTMEKCIKQE-----RFOQLKQNNL 581
 QY 1277 LOQQLVHANKKADNKSKITIDTHF-----LERKMOHH--LLKKEKEIEFNYNHKN 1326
 Db 582 LOQQLDDANKADNQKALINIQARCDARVQNLQAECKRHLLEDNKVLNETHSK 641
 QY 1327 RIYQYEKEKAETE 1339
 Db 642 KECQYEKEKAERE 654

RESULT 10

Q6PF14 ID 06PF14 PRELIMINARY; PRT; 251 AA.
 AC 06PF14; TISSUE=Testis;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smillus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Strausberg R.;
 RL Submitted (Sep-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC057764; AAH57764.1; -
 DR HSP; P20749; IKLB.
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR001969; Pept_Asp_AS.
 DR Pfam; PF00023; Ank; 5.

DR PRINTS; PR01415; ANKYRIN.
 DR SMART; SM00248; ANK; 5.
 DR PROSITE; PS0088; ANK_REPEAT; 3.
 DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
 DR PROSITE; PS0141; ASP_PROTEASE; UNKNOWN_1.
 KW ANK repeat; Hypothetical protein.
 SQ SEQUENCE 251 AA; 27543 MW; 85934A3579801A3A CRC64;
 Query Match 10.8%; Score 750; DB 2; Length 251;
 Best Local Similarity 77.1%; Pred. No. 3.3e-20;
 Matches 148; Conservative 19; Mismatches 25; Indels 0; Gaps 0;
 QY 4 RKKTINLNIODAKRQRTALHWAQVNGHEVTVFLVDRKQDLVDGEGHRTPLMKALQCHOE 63
 Db 59 KKTMDLNRDKARTALYACANGHAEVTVLLVDRKQDLVDGEGHRTPLMKALQCHOE 118
 QY 64 ACANILDSGADINLVYGNMALHYAVYSILSVAKLLSHGAVIEVHNKASLTPLLS 123
 Db 119 ACANILDSGADPNVDVYGNVAVNSVAVKLLSCGADIEVKNKAGHTPLLLA 178
 QY 124 ITKSEQIVFELLTKNANANVYKCTALMLAVCHGSSEIVGMLLQNVDFVFAADICGV 183
 Db 179 IRKSEIEVFELLTKNANANVDFKCTALMLAICHGSSEIVGKLLQNVDFVFAADICGV 238
 QY 184 TAEHYAVTCGFH 195
 Db 239 IAERYAVACGFN 250

RESULT 11

Q60311 ID 060311 PRELIMINARY; PRT; 641 AA.
 AC 060311; TISSUE=Brain;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE KIAA0565 protein (Fragment).
 GN Name=KIAA0565;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=98290545; PubMed=9628581;
 RA Nagase T., Ishikawa K., Miyajima N., Tanaka A., Kotani H., Nomura N.,
 RA Chara O.;
 RT "Prediction of the coding sequences of unidentified human genes. IX.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro."
 RL DNA Res. 5:31-39(1998).
 DR EMBL; AB011137; BAA25491.2; -
 FT NON_TER 1
 SQ SEQUENCE 641 AA; 75777 MW; F11566F2E6A0B9F4 CRC64;
 Query Match 10.8%; Score 745.5; DB 2; Length 641;
 Best Local Similarity 42.0%; Pred. No. 1.2e-19;
 Matches 179; Conservative 69; Mismatches 111; Indels 67; Gaps 6;
 QY 972 RUTLQNEEKRNRADILNEKIREELGRTEOHRKEVKKQLEQALRIODIELKSVESNL 1031
 Db 6 RLAKQNEEKRNRADILNEKIREELGRTEOHRKEVKKQLEQALRIODIELKSVESNL 65
 QY 1032 -----NOVSTHTEHENYLLHEN 1048
 Db 66 DLVQERNDAAQQLSEQDARILQDILTSKQKELMARKKNSEISHRQKEDLFED 125
 QY 1049 CMLKKEIAMLKLEIATLKHOYQKQENKYPEDIKILKKAELQMTLKKEESITKASQY 1108
 Db 126 CMLQEBIALRLLEIDITKQNKQKKYPEDIEAVKKNLQKIKLNEETLTILQY 185
 QY 1109 SGQLKVLIAENTMLTSKLKE-KQDKELBAEIESHHPRLASAVQDHDQIVTSK ----- 1160

```
Db 186 SQGLNLTAEKINLSENGKQNERLBIEMESYRCRLAAAVRDCDSQTDARLKLDFQ 245
QY 1161 KSOEPAFHAGDACLQRKNVDVSTIYNNEVLHOPLSAQRKSKSLKINLVAGDALRE 1220
Db 246 RTRQEWVR-----LHDKMKVDSGLQAKVEILSEKLSNAESKINSLOIQLHNTDRLGR 299
QY 1221 NTLVSEHAQRDQRETQOCQKEAHMYQNEQDNVNNKHTEQOESLDQKLFQLOSKNNWLOQQ 1280
Db 300 ESLILERVQRLDSQTQCKKETEOMYQIBQSLKKYIAQESVEERLSQSQENMLLRQO 359
QY 1281 LVHAKKADNKSII--TTIDHF-----LKKMQHLLKKEKNEIFNYYNNHLKNRIVQ 1330
Db 360 LDDAHKANSQBSTSTIQDQFSAKQLQAESEKQILSQBKNKELMDENYHLKERMDQ 419
QY 1331 YEKEKA 1336
Db 420 CEKEKA 425

RESULT 12
Q81ZM7
ID Q81ZM7 PRELIMINARY; PRT; 453 AA.
AC Q81ZM7;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DE Melanoma-associated antigen (Fragment).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bruno R., d'Orlando O., Altomonte A., Lamaj E., Maio M., Pucillo C.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RE EMBL; AF543495; AAN40505.1; -.
FT NON TER 1
SQ SEQUENCE 453 AA; 53140 MW; 81DF80012122BBCA CRC64;

Query Match 10.6%; Score 736.5; DB 2; Length 453;
Best Local Similarity 42.8%; Pred. No. 1.8e-19;
Matches 172; Conservative 83; Mismatches 120; Indels 27; Gaps 8;

QY 951 KEIKSQLENQKWKVEQELCSVRLTNQBEKEKRNADILNEKIRELGRIEOHRKELEYK 1010
Db 2 EIKSQKHEILLEKELCSLRFATIQEKKKRNVEEVHVKREKLRITEEQYRIEADVT 61
QY 1011 QOQLALRIQDIELKSVESNLNQVSHTEHNYLLHNCMLKKEIAMLKLEIATLKHQYQ 1070
Db 62 KPICKAKSAEVELKTGGNSNQVSETDEKED--LLHENRLMQDELARLEKDTIKNQWL 120
QY 1071 EKENYCFEDIKILKEKNAELQWLTKLKBESLTKRASQVSGQLKVLAENTMITSKL--KEK 1129
Db 121 EK--KYLKDFEIVKRGKHEDLQALKRNGETLAKTIACYSQGAULTDENTILRSKLEQR 178
QY 1130 QDKTELEAIEHSHPRLASVDHDIQVTSRKSQEPAPHIAGDAC--LQRKNVDVSSVI 1187
Db 179 ESRQLETEMQSHYHCLNARNARCDHQSHSSKKRQDELAFQGTVDKRLHQENLNHVL--- 235
QY 1188 YNNEVLHOPLSAQRKSKSLKINLVAGDALRENTLVSEHAQRDQRETQOCQKEAHMYQ 1247
Db 236 ---ILSLQLSKAEKSRVLTGELHYTGEALKEKALVPEHVQSELKQSQMKDIEKMYK 291
QY 1248 NEQDNVNNKHTEQOESLDQKLFQLOSKNNWLOQLVHAKKADNKSIIITIDHF----- 1300
Db 292 SCYNTMEXCIEKOE----RFOCLKQKNMLLOQLDDARKADNQBOKAILNIQARCDARVQ 347
QY 1301 -LERKQWQH--LLKEKNEIFNYYNNHLKNRIVQYEKEKAETE 1339
Db 348 NLQASCRKRLHLLLEEDNKMVLNMLNHSKEKECQYEKEKAERE 389
```

```
RESULT 13
AN18 HUMAN
ID AN18 HUMAN STANDARD; PRT; 992 AA.
AC Q81VF6; Q72468;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DE Ankyrin repeat domain protein 18A.
GN Name=ANKRD18A; Synonym=KIAA2015;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Nagase T., Kikuno R., Ohara O.;
RT "The nucleotide sequence of a long cDNA clone isolated from human.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-349 FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- SIMILARITY: Contains 5 ANK repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB095935; BAC23111.1; ALT_INIT.
DR EMBL; BC056266; AAH56266.1; ALT_TERM.
DR Genew; HGNC:23643; ANKRD18A.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 6.
DR PRINTS; PR01415; ANKYRIN.
DR PROSITE; PS00297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50088; ANK_REPEAT; 4.
KW ANK repeat; Repeat.
FT REPEAT 67 96 ANK 1.
FT REPEAT 100 129 ANK 2.
FT REPEAT 133 162 ANK 3.
FT REPEAT 166 195 ANK 4.
FT REPEAT 199 228 ANK 5.
FT CONFLICT 130 130 K -> E (in Ref. 2).
FT CONFLICT 288 288 Missing (in Ref. 2).
SQ SEQUENCE 992 AA; 115664 MW; B9B288F087340D9A CRC64;

Query Match 9.5%; Score 654; DB 1; Length 992;
Best Local Similarity 18.2%; Pred. No. 3.9e-16;
```

Matches 244; Conservative 121; Mismatches 197; Indels 780; Gaps 13;	
Qy	1 MTKRKTNTLNQDAQKATLHACVNGHEEVTFVLDVDRKQQLDVLGDEHRTPLMKALQC 60
Db	54 LTRFR--DLDRDKRTVLHLACAHGRVQVVTLLHRRQCIDICDRNRTPLMKAVHS 111
Qy	61 HOEACANILISGADINLVGVGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 120
Db	112 QBEACAIVLLECCANPNIKDIYGNTHALHYAVYNGKTSIAERLLSHANIEALNKEGNTPL 171
Qy	121 LLSITKRSQIIVFELLIKNANANAVKYCTALMLAVCHGSSEIVGMLLQONVDVFAADI 180
Db	172 LFAINSRQHWFEFLKNOANIHADVNFRTALILAVOHNLSIVTLLQONIRISSQDM 231
Qy	181 CGVTAHYAVTCGPHHIEQIMEYIRKLSKNHNTNPEGTSGTDPDEAAPLAERTPDFAE 240
Db	232 FQGTAEYALCSDLSIRSQILHKNKMLKHLRNDNQETA-- 273
Qy	241 SLVEKTPDEAALVTPDPAESLVEKTPDEAASLVEGTSDKIQCLEKATSGKPFQSAEE 300
Db	274 ----- 273
Qy	301 TPREITSPAKETSEKFTWPAKGRPKRIAWKEDTTPREIMSPAKETSEKFTWAAKGRPK 360
Db	274 -----MKPA----- 277
Qy	361 IAWKETPVKTCVARVTSNTKVLKGRSKMIACPTKESSTKASANDORPPSKQEE 420
Db	278 -----NLKGRERAKAE-----HNLKVASERKE- 301
Qy	421 DBEYCDLSRSPSSAKIQVCIPESIQVMEINREVEBPKKPSAFKPAIEMQNSVFNK 480
Db	302 ----- 301
Qy	481 APELKNEQTLRADMPPEBKQDYENSWSDESLETSVSKQDVCLPKATHQKEIDKING 540
Db	302 ----- 301
Qy	541 KLEESPNKGLLKATCGMKVSIPTKALELDMQTFKAEPPGKPSAFEPATEMOKSVPNKA 600
Db	302 RLQRSNKQ----- 310
Qy	601 LELKNEQTFRADEILPSESOKDYENSWSDESLETSVSKQDVCLPKAAHQKEIDKING 660
Db	311 -----FQDSQ----- 315
Qy	661 LEGSPVKDGLLKANCCKVSIPTKALELDMQTFKAEPPKPSAFEPATEMOKSVPNKAL 720
Db	316 ----- 315
Qy	721 ELKNEQTLRADEILPSESOKDYENSWSDESLETSVSKQDVCLPKATHQKEIDKINGKL 780
Db	316 -----SYGKKD----- 322
Qy	781 EESPDNDGFLKAPCRMKVSIPTKALELDMQTFKAEPPKPSAFEPATEMOKSVPNKALE 840
Db	323 ----- 322
Qy	841 LKNEQTLRADMPPEBKQDYENSWSDESLETSVSKQDVCLPKATHQKEMDKISGLE 900
Db	323 ----- 322
Qy	901 DSTLSKILDTVHSCERARELQKDHCBQRTGKMEQMKKFCVLKCKLSEAKEIKSQLENQ 960
Db	323 ----- 322
Qy	961 KVKWEQELCSVRLTNQOEKRRNADILNEKIRELGRIBEQHRKELEVKQLEQALRIQ 1020
Db	323 ----- 322
Qy	1021 DIEKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYCKENKYPEDI 1080
Db	323 -----AMYGFMFLKDIAMLKKEIYAIKNDLSRKEKKYIQEI 359

Matches 244; Conservative 121; Mismatches 197; Indels 780; Gaps 13;	
Qy	1081 KILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLAENTMLTSKL-KEKQDKLEILEAEI 1139
Db	360 KSITEINANPEKSVRLNERKITTVARYSOQLNDLKAENARLSELEKEKHNERLEAEV 419
Qy	1140 ESHHPRLASAVQDHQIVTSRKSQEPAFHITAGDACLQKRNVDVSVSTIYNNVHLPSE 1199
Db	420 ESHSLSLATANEYEV-ERKOLELWLRADDSRHEKMGNSISQUTDKNELLTEQVHK 478
Qy	1200 AQRKSKLKNLVAGDALRENTLVSEHAQDQRETOCQKAEHMYQNDNVNKHTEQ 1259
Db	479 ARVKFNTLKGKIRETRDALREKTLALGVSQDLRQAQHRKEMQKHPNGEAKESQSIGK 538
Qy	1260 QESLDOKLFQOSKNMMLQOOLVHAHKADNKSKITIDH--FLERKQVHLLKERNEEI 1317
Db	539 QNSLEERIQOELNELLERQLEDARKEGDK-EIVINIHRDCLENG-KEDLLERNEKEL 596
Qy	1318 FYNHNLKNRIYQYEKEKAETE 1339
Db	597 MKEYNYLKEKLLQCEKEKAERE 618

RESULT 14

Q6S8J5 PRELIMINARY; PRT; 712 AA.

AC Q6S8J5;

DT 05-JUL-2004 (Tremblrel. 27, Created)

DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)

DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)

DE POTE2A (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX PubMed:15276201; DOI=10.1016/j.gene.2004.05.009;

RT Bera T.K., Huynh N., Maeda H., Sathyanarayana B.K., Lee B., Pastan I.;

RT "Five POTE paralogs and their splice variants are expressed in human

RT prostate and encode proteins of different lengths.";

RL Gene 337:45-53 (2004).

DR	EMBL; AY462871; AAS58865.1; -.
DR	HSSP; P20749; IKLB.
DR	InterPro; IPR002110; ANK.
DR	Pfam; PF00023; ANK; 6.
DR	PRINTS; PR01415; ANKYRIN.
DR	SMART; SM00248; ANK; 6.
DR	PROSITE; PS50088; ANK_REPEAT; 4.
DR	PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW	ANK repeat.
FT	NON TER
SQ	SEQUENCE 712 AA; 80752 MW; 12777BACA6AF5137 CRC64;
Query Match 8.6%; Score 594; DB 2; Length 712;	
Best Local Similarity 30.4%; Pred. No. 4.2e-14;	
Matches 172; Conservative 106; Mismatches 218; Indels 70; Gaps 19;	
Qy	9 NLNIQDAQKRTALHACVNGHEEVTFVLDVDRKQQLDVLGDEHRTPLMKALQCHQACANI 68
Db	165 DYNKDKQKRTALHLASANGSEVVKLLDRCOLNLDNKKRTALIKAVQCEDECALM 224
Qy	69 LIDSQADINLVYGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPLLSITKRS 128
Db	225 LLEHGDPNIPDEYGNNTLHYAIYNEDKLMKALLYGADIESKKNHGTPLLLGVHEQK 284
Qy	129 EQIVFELLIKNANANAVKYCTALMLAVCHGSSEIVGMLLQONVDVFAADICGVTAHY 188
Db	285 QQVVKFLIKKANLALDRYGRITALILAVCCSASIVSLLEQNDIVSSQDLSGQTAREY 344
Qy	189 AVTCGFPHHIEQIMEYIRK--LSKNHNTNPEGTSGTDPDEAAPLAERTPDPAESLVEKT 246
Db	345 AVSSHHHVICQLSDYKQKMLKISSENSNPQDLKLTSEEB---SQRFKGSNSQPEKM 401

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:56:15 ; Search time 37.9263 Seconds
(without alignments)
3402.035 Million cell updates/sec

Title: US-09-602-362E-23
Perfect score: 6920
Sequence: 1 MTRKKTINLIQDAQRKA.....NHLKNRIYQYEKKAETENS 1341

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	431.5	6.2	1738	2 T14867	interaptin - slime
2	417.5	6.0	2663	1 S28261	centromere protein
3	408	5.9	3259	1 A56539	giantin - human
4	407	5.9	3225	2 I52300	giantin - human
5	406.5	5.9	3488	2 T34418	hypothetical prote
6	403.5	5.8	1875	2 S38173	myosin-like protei
7	400.5	5.8	2116	2 A26655	myosin heavy chain
8	399.5	5.8	2954	2 T14156	kinesin-related pr
9	393.5	5.7	2253	2 T30336	nuclear/mitotic ap
10	387.5	5.6	1313	2 F96673	hypothetical prote
11	385.5	5.6	1957	2 T38077	hypothetical colle
12	381	5.5	3187	2 J05837	364K Golgi complex
13	377.5	5.5	2139	2 T18296	myosin heavy chain
14	376.5	5.4	5327	2 T13564	microtubule-associ
15	373.5	5.4	2057	2 S61477	myosin II heavy ch
16	368.5	5.3	2017	1 A36014	myosin heavy chain
17	365.5	5.3	1690	2 T13030	microtubule bindin
18	360	5.2	1790	2 S67593	transport protein
19	360	5.2	1964	2 A59282	nonmuscle myosin I
20	358	5.2	1300	2 I53799	CGI protein - huma
21	358	5.2	1410	1 A57013	early endosome ant
22	355	5.1	944	2 S26710	spindle pole body
23	355	5.1	1837	2 T41023	probable nuclear p
24	355	5.1	2168	2 T30171	ninein - mouse
25	353	5.1	2331	2 T25410	hypothetical prote
26	352.5	5.1	1269	2 F94730	probable myosin he
27	350.5	5.1	1935	1 A37102	myosin beta heavy
28	350.5	5.1	2245	2 T18278	myosin heavy chain
29	350	5.1	1356	2 S32763	kinectin 1 - human

30	350	5.1	1938	2 JC5421	smooth muscle myos
31	350	5.1	1972	2 JC5420	smooth muscle myos
32	348.5	5.0	1972	1 A41604	myosin heavy chain
33	348	5.0	1538	2 T29095	cardiac muscle fac
34	347	5.0	1558	2 B71603	RESA-H3 antigen PP
35	347	5.0	1959	1 A33977	myosin heavy chain
36	345	5.0	1940	1 A24922	myosin heavy chain
37	345	5.0	1979	1 S03166	myosin heavy chain
38	344.5	5.0	1961	1 A61231	myosin heavy chain
39	344	5.0	1939	1 A46762	myosin alpha heavy
40	343.5	5.0	1526	2 A45605	mature-parasite-in
41	342.5	4.9	1427	2 S22695	restin - human
42	342.5	4.9	1827	2 T16270	hypothetical prote
43	342.5	4.9	1999	1 S21801	myosin heavy chain
44	341	4.9	1392	2 A43336	microtubule-vesicl
45	340.5	4.9	1940	1 S04090	myosin heavy chain

ALIGNMENTS

RESULT 1

T14867

interaptin - slime mold (Dictyostelium discoideum)

C/Species: Dictyostelium discoideum

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C/Accession: T14867

R/Riveto, F.J.; Kuspa, A.; Brokamp, R.; Matzner, M.; Noegel, A.A.

J. Cell Biol. 142, 735-750, 1998

A/Title: Interaptin, an actin-binding protein of the alpha-actinin superfamily in Dictyo

ts.

A/Reference number: Z18248; MUID:98365468; PMID:9700162

A/Accession: T14867

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1738 <RIV>

A/Cross-references: UNIPROT:O76329; EMBL:AF057019; NID:g3549260; PID:g3549261; PIDN:AAAC

C/Genetics:

A/Gene: abpd

A/Introns: 173/2; 1680/1

Query Match 6.2%; Score 431.5; DB 2; Length 1738;
Best Local Similarity 19.6%; Pred. No. 4.5e-10;
Matches 316; Conservative 270; Mismatches 478; Indels 549; Gaps 70;

Qy	15	AQKXTALHWACVNGHEWVFLVDRK-----COLDVLDDGEHRTPLMKAL-----OCH	61
Db	23	AQKVFVNWCNI-----FLNORSQKIEDLETDLVDGILLGSLLEILSGKNWILSKCK	74
Qy	62	Q-----EACANILDSGADINLVDVY-----GNMALHYAVYSEILSVVAKLLS	104
Db	75	QLKTRLHYINNLFSLKFIGDEG--LRLVGVAESDITDGNLKL-----ILGLVWTLIL	125
Qy	105	HGAVIEVHNKASLTPLLLSTKRSEQIVFLLIKVANANVKNYKCTALMLAVCHGSSEI	164
Db	126	RYQIQSQMNSKSSQONLHSTKPS-----LMLNWKYSQISDYG-----HHIKDL	170
Qy	165	VGMLLQONVDVFAADICGVTAEHYAVTCGPHHIEQIMYIRKLSKXNHTNPEGTSAGT	224
Db	171	TTSP--QNGLLFCALV-----HKLVPEKLDY-KSLG-----RSDSLGN	205
Qy	225	PDEAAPLAERTPTDAESL-----VEKTPDEAAPL--VERTP-----DTAESL-----VEKTP	269
Db	206	LTLAFEVANKELGIPSLDPHDIITTPDELISILYISLFPKVYQOTLEPLNANNNISPSL	265
Qy	270	DEAASLVGEGSDKIQCLEKATSGKFE-QSABETPREITSPAKETSEKFTWAKGRPKRIA	328
Db	266	SSSSSSLLTNPKNKNSIQLSKSTSFEOQNQQOQNLLSP-----NSYRNSIS	313
Qy	329	WEKKEDTPREITMSPAKETSEKFTWAAKGRPKIAWEKKEKTPVKTGCVARVTSNKTVLEK	388
Db	314	FSK-----SPSFEQSQ-----STGSSRSIS--PISSFIKNSTTGNLSKSTSPF-	356

Db	1727	ELERVMEYETLSKKFQSLMSEKDSLSEVQDLKHQIEDNVNSQANLEATEKHDNQNTVT	1786
Qy	731	DE---ILPSESKQDYBESSWDSLSLCTYSQKDVCLPKATHQKIDKINGKLEESPDN-	786
Db	1787	BEGTQSPGETEEQD-----SL--SMSTRPTCSVPSAKSANPAVKDFSSHDEI	1835
Qy	787	DGFLKAPCRMVSIPTKALELMDMTFKAEPPKPSAFAPAEMQKSVNPAKALELKNQET	846
Db	1836	NNYLOQIDQLKRI--AGLE-----BEKQNKKEFSOTLE-----NEKN	1871
Qy	847	LRADQMFPSKQKQVENSWSLSLRTVSQK--DVCVPKATHQKMDKISGLKLESDTS	904
Db	1872	TLLSQISTKDGELKMLQBEVTVMNLLNQIQEELSRTVKLETABEEKDDLEERL--MNQ	1929
Qy	905	LSKILDTVHSCERARELQKDHCEORTG--KMEQMKKKFVCLKKLSKSEAKIISQLENOK	961
Db	1930	LAELNGSI-----GNYCQDVDAQIQEELSRTVKLETABEEKDDLEERL--MNQ	1979
Qy	962	VKWEQELCSVRLTNQEBEK-----RNADIINEKIREELGRIEBOHRKELEVKQOLEAL	1017
Db	1980	TKVESEIRKEYLEKIQGAQKPGNKSHAKELQELLKE-----KQEVKQLQKDCI	2029
Qy	1018	RIQD-----JELKSVESNLQVSHTHENYLLHENCMLKKEIAMKL-----	1060
Db	2030	RYQEKISALERTVKALEFVQTESQ-KOLEITKENLAQAVEHRKKAQAEASFVLLDDTQ	2088
Qy	1061	EIA-----TLKHQYQENKVFEDIKI-LKEKNAELQMTLKLKEESITKASQVSGOLK	1113
Db	2089	SEARVLADNLKLKELQSNK--ESVSKQMKQKDEDLERLEQAEBEKLKKNQKQKLD	2146
Qy	1114	VLIANTMLTSLKE-----KODKEI--LEAIESHPRLA-----SAVQHDQIVT	1158
Db	2147	ALRREKVHLEETIGETVTLNKDKREVQLOENLDSVTVQLAFTKSMSSLODDRVID	2206
Qy	1159	SRKSQEPAPHIAGDACLOKRMVNVSSITYNNEVLHQLSEAQRSKSLKINL-----	1212
Db	2207	EAKKWERK---SDALQSKEEIRLKEB--NCSVLKQOLRQMSIHWEELKINISLEHDK	2261
Qy	1213	--YAGDALRENTL---VSEHAQDQRETOCQMKBAEHMVQNDVNVNKGTEQOESLDOKL	1267
Db	2262	QIWESQAQTEVQLQKQVCDTQGENKELLSQLETRHLYHSSQNELAKLESLKLDQOL	2321
Qy	1268	FQLOSKNMWLOQQ-----LVHAHKADNKSITTDIHFLEKMO-----HLLKE	1312
Db	2322	TDLNSLEKCKEQKNGLEGIHQOQEAQIONKFSYB--QLETDQASRELTSLEHEENM	2379
Qy	1313	KNEEIFNYNHLKNRI-----YQYEKEKAPTEN	1340
Db	2380	KEQKISLLSGKEAIOVAIAELRQHQHKEILEN	2415
RESULT 4			
giattin - human			
N:Alternate names: gcp372			
C:Species: Homo sapiens (man)			
C>Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004			
C/Accession: 152300			
R:Sohta, M.; Misumi, Y.; Fujiwara, T.; Nishioka, M.; Ikehara, Y.			
Biochem. Biophys. Res. Commun. 205, 1399-1408, 1994			
A>Title: Molecular cloning and sequence analysis of a human 372-kDa protein localized in			
A/Reference number: 152300; MUID:95100974; PMID:7802676			
A/Accession: 152300			
A/Status: preliminary; translated from GB/EMBL/DBD			
A/Molecule type: mRNA			
A/Residues: 1-3225 <RES>			
A/Cross-references: UNIPROT:Q14789; GB:D25542; NID:G662389; PIDN:BAA05025.1; PID:9808869			
C/Superfamily: giantin			
Query Match			
Best Local Similarity 20.2%; Pred. No. 9.2e-09;			
Matches 311; Conservative 279; Mismatches 599; Indels 347; Gaps 67;			
5.9%; Score 407; DB 2; Length 3225;			

Qy 962 VKWQELCSVRLTNQBEK-----RRNADIINEKIRELGRIBEOHRKELEVKQOLEAL 1017
Db 1946 TKVESEIRKEYLEIKQAQKPGNKSNAKLELLE-----KQBVKQLQKDCI 1995
Qy 1018 RIQD-----IELKSVEINLQVSHTHENYLLHNCMLKKEIAMKL----- 1060
Db 1996 RYQEKISALERTVVALEFVQTESQ-KDLEITKENLAQAVEHRKKAQAEALASFKVLLDDTQ 2054
Qy 1061 -EIA-----TLKHQVQEKENYFEDIK-LKEKNAELQMTLKLKEESLTKRASVYGOLK 1113
Db 2055 SEAAVLADNLKLKELQSNK--ESVKSQMKQKQEDLERLEQAEEHLKKNQWQEKLD 2112
Qy 1114 VLIENVTMLTSKLE-----KQKEI--LEAEIESHHPLA-----SAVQHDQIVT 1158
Db 2113 ALREKVLHEETIGEIQTNLNKQDEVOQLQENLDSTVTQLAAFTKSMSSLQDDRDRVID 2172
Qy 1159 SRKQOEPAFHITAGDACLQKKNVDVSSITVYNEVLHQLPSAQRKSKLKNLN----- 1212
Db 2173 BAKWERKF-----SDAIQKSEIRLKED--NCVVLKQQLRWMSHMEELKINISRLHDK 2227
Qy 1213 --YAGDALRENTL---VSEHAQRDQRETQCMQKAEAHMYQNEQNVNKHTEBQESLDQKL 1267
Db 2228 QIWESKAQTEVQLQKVCDTLQGENKELLSQLETRHLYHSSQNELAKLESELKSLKQDL 2287
Qy 1268 FOLSKKNMWLQOQ-----LVHAHKADNKSKITIDIHFLERKMQ-----HHLKX 1312
Db 2288 TDLNSLEKCKEQKGNLEGIIRQOEADITQNSKFSYE--QLBTDLQASRELTSRLHEEINM 2345
Qy 1313 KNEEIFYNNHKNRI-----YOYEKEKAETEN 1340
Db 2346 KEQKIISLSGKEBAIQVAIAELRQOQHDKEIKELN 2381

RESULT 5
T34418
hypothetical protein F12F3.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T34418
R:Fulton, B.; Wohldmann, P.
submitted to the EMBL Data Library, July 1998
A:Description: The sequence of C. elegans cosmid F12F3.
A:Reference number: 221521
A:Accession: T34418
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3488 <FUL>
A:Cross-references: EMBL:U80022; PIDN:AAC25985.1; GSPDB:GN000023; CESP:F12F3.3
A:Experimental source: Strain Bristol N2; clone F12F3
C:Genetics:
A:Gene: CESP:F12F3.3
A:Map position: 5
A:Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match 5.9%; Score 406.5; DB 2; Length 3488;
Best Local Similarity 21.6%; Pred. No. 1.1e-08;
Matches 273; Conservative 217; Mismatches 519; Indels 253; Gaps 48;

Qy 213 QNTNPEGTSAGTPDEAALPABRTPDABSLVEKTPDEAALPVER-----TPDTAESL 264
Db 311 QAINVAGEAWCFSDVVMHMESESRDDDKSVDE--VDDSTVLVEEKDDGDDKSKPKTKKKI 368
Qy 265 VEK--TPD-EAASLVEGTSKDIQCLEKATSGKFSQSAETPREITSAPKSETKFTWPAK 321
Db 369 IKKETPESEQVTAABPEQKIKSEVD-----VQSVAE--EVGAKKPPDAEKPTDLISK 419
Qy 322 GRPRKIAWEKKEDTPREIMSPAKETSEKFT---WAAKGRPRKIAWEKKETPVKTCGVARV 378
Db 420 A--KQDSKSKSDP-EASTEESTTEKPTNDKTSKSAEKKTVKPKKEVTGKPLEAKKP 476
Qy 379 TSNKTKVLEKRSKMIACPT-----KESSTKA-----SANDQRPSESQOEDEEYSCDSRS 430
Db 477 VEDKDDASQPSSESSPPTDGKKKQIPKALFIPDEISSRFGDPSTHSHSETNITITIRG 536

Qy 431 LFESSAKIQVCIPE-----SIYQKVMENRVEEPPKPSAFKPAIE-----MNSVYP 478
Db 537 R-EGSADAKTPLVPLSASVSMKVFTLVESAKE--KAEFSFKRRSETPDDKSRKKEGLPP 593
Qy 479 NKAFLKNEQTLR---ADMPFPPEKQKDYENSWDSSELCTVTSQKDVCLPKATHQKEI 535
Db 594 AKSEKEDVTAEBKQSTEALI--ESKKEVDSEKISEQQPSDDKNKSEVVGPEKAAGPET 651
Qy 536 DKINGKLEESPNKQGLLKATCGMKVSIPTKALELKMOTFKAEPPGPKPSAFEPATEMOKS 595
Db 652 KQDVSEIEVPKKTIKKTEKSDSSISQSNVLKPADDKSKSDVDVTDKSKKTTEDQTK 711
Qy 596 VPKALELKNQOTWR---ADBIPLSESKQDYENSWDTESL-----CE 636
Db 712 VATDSKLEKAADTTKQIETETVDDKSKKVLKKTKEKSDSFISQKSETPPVVPEBTPKAE 771
Qy 637 TVSQKDVCLPKAAHQKEID---KINGKLEGSFVKDGLLK--ANCGMKVSIPTKALELMDM 691
Db 772 SEAQKIAEVNKAQKQKEVDNLRKBAEVAAKIADEKLKIEAEANIKKTAEVEAAK---- 827
Qy 692 QTFKAEPEKPSAFEPALTEMOKSVPNKALELKNQOTLRADBIPLSESKQDYEESSWDSE 751
Db 828 ---KQKEDQOLKLETVVSKSAEK--LELEKQAIKKAABADAVKKQKELNEKV----- 879
Qy 752 SLCTVSKQVCLPKATHQKEIDKINGKLEESPNDGFLKAPCRMKVS-----IPT 802
Db 880 -----KLEAAKSAADKL--KLEEE-----SAAKSKVSESVKFGEBKT 918
Qy 803 KALELMDQTPKADBPPEKPS-----AFEPALTEMOKSVPNKALELKNQOTLRADQMPFS 855
Db 919 KAGE--KTVQVESEPTSKKTTIDTVGATEPADETPKKIIKKGTEKSDSSISQKSAADS 976
Qy 856 E--SKQKVEENSWDSLSRETQVSKQVCPVKATHQKEMD-----KISGKLESDTSL 906
Db 977 EKVSKQKEQDEPTPAVSETQMTVEAD---KSKQKETDSEKLDLDAEIAKTKQEADEK 1032
Qy 907 KILDTVHSCERARELQKDHCEQRTGKM-----EQMKKFCVLKKLKSAKBIKSQLE 958
Db 1033 SKLD--AQEKIKKVSDDAARKEKELNDKLKSEIATKASADKLKLEQAQAKAAE 1089
Qy 959 NQKVWQELCSVRLTNQBEKERNADILNEKIREELGRTEBQ-----HRK 1005
Db 1090 VEAARKQKE-----KDEQLKLDTEAASKKAAAEKLELEKQAIKKAAGADAVKKQK 1140
Qy 1006 ELEVKQOLEQALRIQDIELKSVEINLQVSHTHENYLLHNCMLKKEIAMLKLLEIATL 1065
Db 1141 ELDEKNKLEANKKSAAGKLKIEESAAKSKQTVBEO-----AKLDAQT- 1183
Qy 1066 KHQYQEKENYFEDIKILKEKNABLOMTLKLKESLTKRA-----SOYSGQLKVLJ-- 1116
Db 1184 KAKTAEQTKLEKDEKSTKSESEKTVDEPKPKVVKKTEKSDSSISQKSETSKTVVES 1243
Qy 1117 -----AENTMLTS---KLKEQKQDEKILEAEIESHHPLRSLASAVQHDQIVTSRKSEPAFH 1168
Db 1244 AGPSESETQKVADAAKQKETDEKQLEAEI-----TAKSADSEKSK 1285
Qy 1169 IAGDACLQKKNVDVS--STIYNEVLHQLPSAQRKSKLKNLNAG-----DALR 1219
Db 1286 LEAESKLUKKAEEVAARKQKEKDEQLKLDTEAASKKAAAEKLELEKQSHIKKAABEVDVK 1345
Qy 1220 ENTLVSHAQRDQRETQCMQKAEAHMYQNEQNVNKHTEBQESLQKLFQLSQKNMWLQO 1279
Db 1346 KQKEL--BEKQLESEAAATKADAELKLEEQ-----KKKAAELALIEIOKE---QE 1392
Qy 1280 QLVHAHKADNKSKITIDIHFLERKMOHLLKKNBEIFNTNNHNLKRNRIYQVEKEKATE 1339
Db 1393 KLAQEQSRLEDEAKSAEKQKLESETSKQTEEPAKB--SVDEKPKKVLKKTKEKSDSS 1450
Qy 1340 NS 1341
Db 1451 IS 1452

RESULT 6

S38173
 myosin-like protein MLP1 - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein YKR095w; protein YKR415
 C:Species: Saccharomyces cerevisiae
 C>Date: 03-May-1994 sequence revision 03-May-1994 #text_change 09-Jul-2004
 C:Accession: S38173; S40647; S31207
 R:Baladron, V.; Ballesta, J.P.G.; Bou, G.; del Rey, F.; Esteban, P.F.; Garcia-Cantalejo, J.
 submitted to the Protein Sequence Database, March 1994
 A:Reference number: S38158
 A:Accession: S38173
 A:Molecule type: DNA
 A:Residues: 1-1875 <BAL>
 A:Cross-references: UNIPROT:Q02455; EMBL:Z28320; NID:G486586; PID:G486587; MIPS:YKR095w
 A:Experimental source: strain S288C
 R:Bou, G.; Esteban, P.F.; Baladron, V.; Gonzalez, G.A.; Cantalejo, J.G.; Remacha, M.; Jilka, J.
 Yeast 9, 1349-1354, 1993
 A:Title: The complete sequence of a 15 820 bp segment of Saccharomyces cerevisiae chromo-
 some V
 A:Reference number: S40644; MUID:94205265; PMID:8154186
 A:Accession: S40647
 A:Molecule type: DNA
 A:Residues: 1-1875 <BOU>
 A:Cross-references: EMBL:X73541; NID:G450550; PIDN:CAA51948.1; PID:G450554
 A:Experimental source: strain S288C
 R:Kobling, R.; Nguyen, T.; Chen, E.Y.; Botstein, D.
 Mol. Gen. Genet. 237, 359-369, 1993
 A:Title: A new yeast gene with a myosin-like heptad repeat structure.
 A:Reference number: S31207; MUID:93247549; PMID:8483450
 A:Accession: S31207
 A:Molecule type: DNA
 A:Residues: 1-300, 'A', 302-1875 <KOE>
 A:Cross-references: EMBL:L01992; NID:G171958; PIDN:AAA34783.1; PID:G171959
 C:Genetics:
 A:Gene: SGD:MLP1
 A:Cross-references: SGD:S0001803; MIPS:YKR095w
 A:Map position: 11R

Query Match 5.8%; Score 403.5; DB 2; Length 1875;
 Best Local Similarity 21.2%; Pred. No. 6.7e-09;
 Matches 334; Conservative 258; Mismatches 527; Indels 453; Gaps 73;

QY 40 KQQLDVLGCHRTPLMKALQCHQACANILDSGADINLDVYGNMALHYAVISILSV 99
 DB 124 KKKLDDLTETEE-----KKTQSNQOQTLKILDERLKEIELVVENNRS-----NSEC 169

QY 100 AKLSHGAVIEVHNKASITPLLSITKRSQEQIVFLLIKNAN----- 141
 DB 170 KKLSTIMDLTKQOQYITNDLSNRTELEKTKOELTLQSNNDWLEKELRSKNEQYLSYR 229

QY 142 -----ANAVNKYKCTALMLAVCHGSSIVGMLLQONVDVFAADICGVTAEHYAVT 191
 DB 230 QKTDKVIDIRNELRLNDFOMERTNN-----DVLKQKNNELSKS----- 270

QY 192 CGFHHIHEQIMEYIRKSKNHNTPGTSAGTDPDEAAPLAERTPDTAESLVEKTPDRAA 251
 DB 271 -----LOEKLE-IGKLS-DSLSEKQEFSA-----ENSLKQRLVDLLESQNLNAVEELN 318

QY 252 PLVETPTDPAESLVE-----KTPDEAASLVEGTSKIQCLEKATSKFPQSABETPR--EI 305
 DB 319 SI--RELNTAKVIADDSKKQTPENEDLLKE-----LQTLKEKLAQCEKECLRLSSI 367

QY 306 TSPAKETSEKFTWPAKGRPRKTAWEKEDTTPREIMSPAKETSEKFTWPAKGRPRKIAWEK 365
 DB 368 TDEADEDNENLS--AKSSSDFTPLKKQLIKERTKEHLNQOITETFIVELEHKVPIINSFK 425

QY 366 KETPV-----KTGCVARTSN-----KTKVLEKGRSKMIACP-----TKES----- 401
 DB 426 ERTDMLNELNNAALLHETSNEKNAKVKELNKQKLVCECENDLQTLTKQRLDLCRQIQ 485

QY 402 ----STKASANDORFPSESQ-----EEDREYSCDSRLSFESSAKIOVCIPESYQKV 450
 DB 486 YLLITNSVNSDKGPLRKEIQFIQNIQMEDDS-----TITESDSQKV 528

QY 451 MEINREVEEPPKPSAFKPAIEMQNSVPNKAFELKNEQTLR-----ADPMFPPEKQKOY 505
 DB 529 V-TERLVE-----FKNIIOQE-----KNAELLKVVRNLADKL---ESREKKS 567

QY 506 EENSWSDESICETVSQ-KDVCL-----PKATHQKEIDKINGKLEE-----SPKQGLLKATC 556
 DB 568 KQSLQKIES--ETVNEAKEAIIITLKEGMDLESRIEELQKELBELKTSVFNEDASYNVT 625

QY 557 GMKVSPTKALE--LKDQTFKAEPPKPSAFEPATEMOKSVPNK-----ALE 602
 DB 626 IKQLTETKRDLESQVQLQF-----RISQITRESTENSLNKEIQDIYDSKDISIK 678

QY 603 LKNEQTRWADEILPSE-----SKQKDYEN-----SWTESICET 637
 DB 679 LGREKESR---ILAEERFKLLSNTLDITKAENDQLRKRFDYLTQNTILKQDSKTHETLNEY 735

QY 638 VSQK---DVLCPKAAHQEIDKINGKLEG-----SPKQGLLKANCCKMVISITKA 685
 DB 736 VSCSKLSIVETELLNLKERQKLRVHLEKNLKOELNKLSPKDSL-----R 781

QY 686 LELMDQTFKAEPPKPSAFEPATEMOKSVPNKALELKNQTLRADEILPSESQKDY-- 743
 DB 782 IMVTQLQTLQ---KEREDLLE---ETKSCQKKIDELED-----ALSELKETSQKDHII 830

QY 744 ---EESS---W-----DSESLCETVSQKOVCLPKATH-----QKEI--DKIN- 777
 DB 831 KQLEEDNNSNIEWQNKIEALKXDYESVITSVDKQTDIEKLVKVSLEKEIEEDKIRL 890

QY 778 ---GKLEESPDNGFLKAPCRMKVSIIP---TKALELMDM-----OTFKASPPKPSAFEP 826
 DB 891 HTYNVMDETINDSLRKELEKSKINLTDAYSQIKYKDYLTETTSQSLQQTNSKLDESFKD 950

QY 827 AIENQKSVPNKALELKNQTLRADQMFPPSESK---QKKVEENSWDSLSLETVSQ---KD 880
 DB 951 FTNQIKNLTKETSLEDKISILLKEQMFNLNELLQKGMKEKADFKKISILQNNKE 1010

QY 881 VCVPKATHQKEMDKISKLEBDST-----KRNADIILNEKI----- 992
 DB 1011 VEAVKSEVESKLSKIQNDLQOQTIYANTQNNYEQELQKHADVSKTISLSEQLHTYKQG 1070

QY 917 -RARELQKDHCERTQMGQMKKFCVLLKKLSAKEL--KSQLEN---QKVKEQ--- 966
 DB 1071 VKTLNLSRDQLE---NALKENEKSWSQKESLLEQLDLSNRIEDLSQNKLLYDQIQY 1127

QY 967 -----ELCSVRLTLNQEES-----KRNADIILNEKI----- 992
 DB 1128 TAADKEVNNSTNGPLANNILITLRERDILDTKVTVAEKDAKMLRQKISLMDVBLQDART 1187

QY 993 -----REELGRIBEQHRKELEVKQLEQALRIQDIELKSVESNLSNOVSHTHENYL 1044
 DB 1188 KLDNSRVEKENHSHIIQCHDDIMEKLNQL-NLLRESNITLR-----NELENNNNKKEL 1240

QY 1045 LHENCMKKETAMLKLEIATLKHQYQSKENKYFEDIKILKEKNAELQMTLKLKESLTKR 1104
 DB 1241 QSEYDLKQNVAPIESELTKYSNQKE-----QELKLAKE---EVHWRKRSQDILEKH 1293

QY 1105 ASQVSGQVLIAENTMLTSLKLEK-QDKETLEAEIESHHPRLASAVQDHQIVTSRKSQ 1163
 DB 1294 EQLSSDYKLESE-----IENLKELENKERQGAEEKFNLRRAQ--ERLKTSLKSQ 1347

QY 1164 EPAPFHAGDACLQKRMNVDSSTIYNNVHLHQPISEAQRKSKSLKINLNYAG-----DAL 1218
 DB 1348 D-----SLTEQVNSLRDAKNVLENSLSSEANARIEELQ-NAKVAQGNQLEAI 1393

QY 1219 RENTLVSEHAQRDQRETCOMKEAHMYQNEQDNN-----KHTEQESLQDKLQFOLQ 1271
 DB 1394 RK---LQEDAKAGRELQAKLEESTSYESTINGLNEEITTLKKEIEKQROIQQOQOATS 1450

QY 1272 SKNMWMLQOOLVHAHKADNKSKITDIFLERKQKHLLKKEKNEEIPFYNNHL---KNRI 1328
 DB 1451 ANEQNDLSNIVESMKKSFEEDK---IKFIKEKTQ-----EVNEKILEAQERLNQPSNIN 1501

Qy 1329 YOYEKKAETEN 1340
Db 1502 MEIIRKKWSEH 1513

RESULT 7
myosin heavy chain [similarity] - slime mold (Dictyostelium discoideum)
N;Contains: myosin ATPase (EC 3.6.4.1)
C;Species: Dictyostelium discoideum
C;Date: 05-Oct-1998 #sequence revision 05-Oct-1998 #text_change 09-Jul-2004
C;Accession: A26655; A24728; S00250
R;Warrick, H.M.; De Lozanne, A.; Leinwand, L.A.; Spudich, J.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 9433-9437, 1986
A;Title: Conserved protein domains in a myosin heavy chain gene from Dictyostelium discoideum
A;Reference number: A26655; MUID:87092266; PMID:3540939
A;Accession: A26655
A;Molecule type: DNA
A;Residues: 1-2116 <WAR>
A;Cross-references: UNIPROT:P08799; GB:M14628; GB:M11938; NID:G167834; PIDN:AAA33227.1;
R;DeLozanne, A.; Lewis, M.; Spudich, J.A.; Leinwand, L.A.
Proc. Natl. Acad. Sci. U.S.A. 82, 6807-6810, 1985
A;Reference number: A24728; MUID:86016788; PMID:3901008
A;Accession: A24728
A;Molecule type: mRNA
A;Residues: 2035-2116
R;Wagle, G.; Noegel, A.; Scheel, J.; Gerisch, G.
FEBS Lett. 227, 71-75, 1988
A;Title: Phosphorylation of threonine residues on cloned fragments of the Dictyostelium myosin heavy chain
A;Reference number: S00250; MUID:88112226; PMID:2828113
A;Accession: S00250
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1734-1893 <WAG>
C;Comment: The rod domain is highly periodic, containing a pattern of 7-residue repeats
C;Superfamily: myosin heavy chain; myosin motor domain homology
C;Keywords: actin binding; ATP; coiled coil; hydrolase; nucleotide binding; P-loop; phosphatase
F;1-818/Domain: globular head <HED>
F;89-747/Domain: myosin motor domain homology <MMOT>
F;179-186/Region: nucleotide-binding motif A (P-loop)
F;819-2116/Domain: alpha-helical rod <ROD>

Query Match 5.8%; Score 400.5; DB 2; Length 2116;
Best Local Similarity 20.2%; Pred. No. 1e-08;
Matches 266; Conservative 253; Mismatches 470; Indels 329; Gaps 55;

Qy 226 DEAPLAERTPTAESLVEKTPDEANPLVE--RTPTDAESLVEKTPDEAASLV---EGTS 280
Db 908 DEKKLALENQKRSVEEKVRDLEELQEEQKLRNTLEKLUKKYBELEEMKRVNDGQS 967

Qy 281 DKIQLEKATSGKPEQSAEETPREITSPAKETSEKFTWPAKG-----RPRKIAWEKKEDT 335
Db 968 DTISLEKID-----ELQREVELTESFSEESKDKGVLEKTRVRLQSELDL 1015

Qy 336 PRETMSPAKETSEKPTWAAGRPKIAWEKK-----ETPVKTCGVARTSNK----- 382
Db 1016 TVRLDSETKDSKSELRL-----QKKLEELKQVQEAALAAATAAK---LAQEAANKKLQGE 1067

Qy 383 -TKVLEKGRSVMACPTKESSTK-----ASANDORPSESQBED----- 421
Db 1068 YTELNEKFNSEVTARSNNVEKSKTLESQLVANNEL--DEBKNRDALEKKKKKALDAMLE 1125

Qy 422 -----EYSCDSRLSFESSAKIOVCIPESYQKVMENREVEPPPKKPSAFKPAIEMQN 475
Db 1126 EMKQLESTGGEKSLVDLVKVK-QESDMEALRNQISELQSTIAKLEIKST-----LEG 1178

Qy 476 SVPNKAFELNKEQTLRADMPPPPSKQKDYENSWDSLSLCTETVSQKDVCLPKATHOKEI 535
Db 1179 EVARLQGELEAEQAKSNVQKQKKVLELDLDDK---SAQLAEETAAK-----QAL 1225

Qy 536 DKINGKLEESPNKQGLLKATCGMKVSPITKALELKDQMTFKAEPGPKPSAFEPATE-MQK 594
Db 1226 DKLKKKLEQ-----ELSEVQTLSEANNKNNVNSDSTNKHLET 1262

Qy 595 SVPNKALELKEQETWR-----ADEILPSESKQD-YEENSMDTESLCEET 637
Db 1263 SFNNLKLELEAEQAKQKALEKRLGLESELKHVNEQLEEEKKQKESNEKRVLDLEK--EV 1320

Qy 638 VSQKDVCLPKAAHOKETDKINGKLEGGPVKGLKANGCMKVSPITKALELMDMTQTFKAE 697
Db 1321 SELKDQIEEVEAVSKAVTEAKNKE-SELDS--IKROYADVSSRDKSVE--QLKTQAK 1375

Qy 698 PPEKPSAFEP-----TEMOKSVNKALELKEQETLRADEILPSSKQD 742
Db 1376 NEELRNTAEAEAGQLDRAERSKKAEPDLEAVKN--LEETAKKVKAAKAM--KKASTD 1431

Qy 743 YEESWDSLSLCTETVSQKDVCLPKATHOKETDKINGKLEESPDN-DGFLKAPCRMKVSP 801
Db 1432 YRSTKSELDDAKNVSSQYVQIKEL--NEELSELRSVLSEADERCNSAIKAKKTAESALE 1489

Qy 802 TKALELMDMTQTFKAEPEKPSAFPAI-EMOKSVNKALELKE-----QTLRADQ 851
Db 1490 SLKDEIDAAANNAKAKAERKSKLEVRVAELSELESKGTYNVFIRKDAEIODDLARL 1549

Qy 852 MFPSESKQKYEENSWDSLSLRETVSQKDVCPKATHQKE--MDKISGKLE-DSTSLSK 907
Db 1550 DRETESRIKSD-----DKNTRKQFADLEAKVEEA--QREVVTIDRLKKLESIIIDUST 1603

Qy 908 ILDT-----VHSCERARELQKHCEQRTGM-----EQMKKRFK----- 941
Db 1604 QLDTEKTSRIKIEKSKKLEQTLAERRAAEBGSKAADEBIRKQVQWQVDSLRADLDSER 1663

Qy 942 -----VLKKLSEAEKISQLENO-----KVKWEQELCSVRLTLNOBEEKQR 983
Db 1664 AALNASEKKIKSLVAEVDVEKQLEDEILAKDKLVKAKRALEVELEVRDQLEBEDSRS 1723

Qy 984 NADILNEKIREELGRIBEOHRKELEVKQOLEQALRIQDIELKSVESNLNOVSHTHENNY 1043
Db 1724 ELEDKRLITTEVEDIKKYDAEVEQNTKLEA-----KKKLTDDVDTLUKKQLEDEKK 1776

Qy 1044 LLHENCMKLKEIA-----MLKLEIATLKHQYQEKENKYPF-DIKILKEKNAELQMTLKL 1096
Db 1777 KLNESERAKKLESENEDFLAKDAEVKNRSRAEKDRKKYKEDLKDTK-----YKL 1827

Qy 1097 KEESLTKRASQY-SGQLKVLIAENTMLTSLKKEVQ-----DKLEIAEIESHHPRLA 1147
Db 1828 NDEAAKTQTEIGAALKEDQIDE---LRSLKEQQAQATQADKSKKTLEGSIDN----LR 1880

Qy 1148 SAVODHQIIVTSRKSQSPAFHIAQDACLQRMNVDSSTIYNNVHLQPLSEAOQKSKSL 1207
Db 1881 AQIEDEGKIKMRLSEKRA-----LEGELE-ELRETV---EEAEDSKSEASQSKSLV 1928

Qy 1208 KINL-----NYAGDA-----LRENTLVSEHAORDORETOCOM 1239
Db 1929 ELELEDARRNLQKEIDAKEIAEDAKSNLQREIVEAKGRLEBESIARTNSDRSKRLEAEI 1988

Qy 1240 K-----EAHMYQNEQDNVNHKTEQESLQKLF-----QLQSKMMVLOQLVHAHKA- 1288
Db 1989 DALTAQVDAQAKNQOIKENKKTIELKEYRKFGSEKTKTKFEFLVEKLETDYKGA 2048

Qy 1289 -----DNKSKITIDIHFLERQMOHLLKEKNEEIFNYYNNHLKNRIYQVEKEKAETENS 1341
Db 2049 KEAADEQQRLTVENDLRKHLSEISLLKDAIDKL-----QRDHDYKKELETETA 2098

RESULT 8
T14156
kinesin-related protein - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T14156
R;Wood, K.W.; Sakowicz, R.; Goldstein, L.S.; Cleveland, D.W.
Cell 91, 357-366, 1997
A;Title: CNP-B is a plus end-directed kinetochore motor required for metaphase chromosome
A;Reference number: 217893; MUID:98028574; PMID:9363944
A;Accession: T14156

Db 473 YEKMLEDARNEIDSLKS-----TVDSIQNEFENS KAGWEQKELHLMGCVKKBEE-----NS 524

Qy 530 THOKIDKINGKLESPNKDGLLK--ATCGMKVSIPTKALELKMOTKAPPPGKPSAF- 586

Db 525 SSQEEVSLVNLKES--BEDACARKEEASLKNLKVAGGEYKYLQETLGEAKAESMKLK 583

Qy 587 EPATWOKVSPNKALELNEQOTWRADEILPSPSKOKDYEEENSWDTESELCTVSKQDVCLP 646

Db 584 ESLDKEDLKNVTAIEISLREWEGSVLEKIEELSK-VKESLVDKETKLOISITQABEELK 642

Qy 647 --KAARQKEIDKI-----NGKL-----EGSPVKD-----GLLKANGCMKVSIPTKAL 686

Db 643 GREAAHMQIEBELSTANASLVDATKLIQSVQESDLKEAGYLLKTEELSVANESLAD 702

Qy 687 ELMDMOTKAP--PEKPSAFAPAEIMOKVSPNKALELNEQOTL-----ADELPSB- 737

Db 703 NYTDLQSVQESKOLKEREVAYLKIE-ELSVANESL-VDKETKLOHIDQEAELRGREA 760

Qy 738 SKQKDYESSWDSSELCTVS-----QKDVCLPKATHOKEIDKI--NGKLBPSPDN 786

Db 761 SHLKKIEBELKENENLVNANVMQNAIESKDLREREVAYLKIDELSTANGTLDADVTN 820

Qy 787 DGLFKAPCRMKVSIPTKALELMDMOTKAPPEKPSAFAPAEIMOKVSPNKALEL----- 841

Db 821 -----LQNISEENKELRERETLLKKAELS-----ELNESLVDKASKLQTVVQ 864

Qy 842 KNEQTLRADOMPSPSKOKKYEENSWDSSELCTVSKQDVCPKATHOKE--MDKISGKL 899

Db 865 ENEELRERETAY-----LKKIEELS-----KLHEILSDQETKLIQSNHEKEELKERETAYL 915

Qy 900 EDSTLSKILDTVHSCERARELQDHCQORTCKMEQKKFCVLKKKLSAEKESQK-- 957

Db 916 KIEBELSKVQEDLLN-----KENELHGMVVEIEDLRSKOSLAQKIEELSNFNVASLLI 968

Qy 958 -----ENQKWKVEQ-----ELCSVRLTLNQEERKRNADILNEKIRBE----- 995

Db 969 KENELQAVVCENEELSKSQVSTLKTIDELSLDKQSLIHKEKELOAAIVENKELKAEALS 1028

Qy 996 LGRIEQRKHELVKQOLEALRIODIELKSVESNLQVSHTHENENYLLHNCMLKKEI 1055

Db 1029 LQRIEE-----LTLNKQTLIDKONELQV-----PHNEELKAKEBASLLKID 1071

Qy 1056 AMLKLEIATLKHQYQENKYPEDIKILKEKNAELQMTLKEESLITKRASQYSQKVL 1115

Db 1072 ELLHLE-----QSWLEKSEF-----QRTQENLELKTQDALAAKIEELSKLES 1118

Qy 1116 TAENMTLSKLEKQDKTEILEAEIESHPRLASAVQDHDQIVTSRKSOEPAPHIAGACL 1175

Db 1119 LEKETEL--KCREAAALEKMEBPSPKHGNSLSICKQYDLVQFS-----EVNGASNG 1168

Qy 1176 QKQMVVDVSTIYNNEVLHPLUSEAQRKSKSLKINLNTAGDALRENTLVSEHAQDQRET 1235

Db 1169 DEKTKTDHYQORSRHEMTIQESPEAIDK-----HLMGERRAITHKVAHRVEGER 1216

Qy 1236 QOMQKEAE-----HMYQNEQDNVNHKTEQ-----QBSLDQKLPOLQSKNNWLOQQLVHAHK 1286

Db 1217 NVE-KESFQKWDYSYKIEKSEVSPRETELDSVEBEVDSK--AESSENW---DQYNGFS 1270

Qy 1287 KADNKSKITIDTHFLERQWQHLLKKEKEEIEFNYNHLLKNRIYQYKEK 1335

Db 1271 LTDDH-----TEDSGNLLLEQH--MKKKPLLRKFGNLLKKSTSSSQK 1313

RESULT 11

T38077
 Hypothetical coiled-coil protein - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: T38077
 R:Connor, R.; Churcher, C.M.; Barrall, B.G.; Rajandream, M.A.; Walsh, S.V.
 submitted to the EMBL Data Library, April 1996
 A:Reference number: Z21767
 A:Accession: T38077

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1957 <CON>

A:Cross-references: UNIPROT:Q10411; EMBL:Z70690; PIDN:CAA94624.1; GSPDB:GN00066; SPDB:SP

A:Experimental source: strain 972h-; cosmid c1f3

C:Genetics:

A:Gene: SPDB:SPACIF3.06c

A:Map position: 1

Query Match 5.6%; Score 385.5; DB 2; Length 1957;
 Best Local Similarity 19.4%; Pred. No. 3.8e-08;
 Matches 292; Conservative 270; Mismatches 574; Indels 369; Gaps 61;

Qy 71 DSGADINLVYVGNMHLHYVSEILSVVAKLLSHGAVIEVHNKASLTPLLLSITKRESEQ 130

Db 62 DSSSELKQKVRG-MRRH-----SDLSIDAKLGS-----SEGSTASSALPLTPRSPS 107

Qy 131 IVEFLLIKNA-----NANAVNKYKCTALMLAVCHGSSEISVGMLLQON--VD 174

Db 108 NASWLLVRGGLDSPILDINSVTQKSNLLNELKQVRSKLAALHEHENGILSLQSSSKKD 167

Qy 175 VPAADICGVTAHVAVTCGFHHIHOIMYIRKLSKNHNTNPEGTSAGTDPDEAPLAER 234

Db 168 KNTSSVTLTSE-----EDVSFQKKL-----TNMESNFAKQSEAYDLGRQ 209

Qy 235 TPDTAESLVEKTPDEAAPLVERTPDTAESLVEKTPDEAAS--LVEGTSDKIQCLEKATSG 292

Db 210 LLTVTEKLDKKEKD-----YEKIKEDVSIKASLAEOQSNKSLRGEORLEKL-LVSN 263

Qy 293 KPEQSAEETPREITSPAKETSEKTPWA-KGRPKRIAWKK-----ED 334

Db 264 KTVSTLRQTENSLRAECKTLQEKLEKCAINEEDSKLEELKHNVANYSDAIVHKDKLIED 323

Qy 335 TPREIMSPAKETSEKFTWAAKGRPKRIAWKKE-----TPVKTGCVAR 377

Db 324 LSTRISEFDNLKSERDTLSIKN-----EKLEKLLRNTIGSLKDSRTSNQSEEMVEL 376

Qy 378 VTSNKT--KVLKGRSKMIACPTKESSTKASANDORPPSEKQBEDEYSCDSRSLFPSS 435

Db 377 KESNRTHTSOLTDAESKLSFEQENKSLKSIDEVQNNLSKDKVMKVQVSQLEARS 436

Qy 436 AKIQVCIPEISYQKWEINREVBPPKPSAF-KPAIEMQNSVPNKAFELKNEQTLRADP 494

Db 437 AHAT-----GKLAEINSEKDFQNKIKDFEKIEQDLRACLNLSNSELKESAL---- 484

Qy 495 MPPESSKOKDYEEENSWD-----SESLCTVSK--QKDVCLPKATH--QKEIDKING 540

Db 485 -----IDKKQDELNNLREQIKERQKVSSESTOSSQSLQORDILNEKKHVEYVESQNLK 539

Qy 541 KL-----EESPNKDGILLKATCGMKVSIPTKALELKD 571

Db 540 ELQTEISNSEHLSQSLTLAAEKAATAVNNELSESKNLSQTLNFAFEKLAQSVMLK 599

Qy 572 -MOTPKABPPGKPSAFAPAEIMOKVSPN-----KALELKNEQOTWRADEI 614

Db 600 NEQNFSS-----LDTSFKKLNEHQLENNHOTITTKQLKDTSSKLQOLQLEAFQKEST 655

Qy 615 LPSES-----KQDYEEENSWDTESELCTVSKQDVCLP-----KAAHQ 651

Db 656 LSDENNDRLTLLKLEESNKSLLIKQEDVDLSLEKNIQTLKEDLRKSEEARFKLEAKNL 715

Qy 652 KE-IDKINGKLE-----GSPVKDGLLKANGCMKVSIPTKALE-----LMDMQTPKA 696

Db 716 REVIDNLKGGHETLEAQRNDLHSSLSDA--KNTNAILSELSTKSSDEKVLTAHVTLTQ 773

Qy 697 EPPEKPSAFAPAEIMOKVSPNKALELKNEQOT---LRADEIILPSESQKDYEEENSWDSSL 753

Db 774 DSKAMQKQFTSLVNSYQSIISNLXHELDRDHVMQSNNTLLESESKLK-----TD 823

Qy 754 CETVSKQDVCL-----PKATHQ-----KEIDKINGKLEESPDNDGFLKAPCRMKVS-- 799

Db 824 CENLTQQNMTLIDNVQKLMHMHVQESKVSSELKEVNGKLSLDLKN---LRSLNVAISDN 880

QY 800 --IPTKALEL-MDMOTFFAEPPEKSAFEPALIMOKSVPNKALELKNFOT-LRADQMPPS 855
Db 881 DOILTQLAELSKNYDSLEQSAQLNGLK-SLEAEKQL-----LHTNEBELHRLDKL--- 932
QY 856 ESKQKKVEENSWDSLESTETVSQKDVCPKATHQKEMDKISGKLEDDSTLSKILDTVHSC 915
Db 933 -TGKLIKIESKSSDLGKLTARQEBISNLKEENWSQSAI-----TSVSKSLDETUSK 984
QY 916 ERARELQKHCEQRTGKMEQMKKFCVLKKKLSEAKEIKSQLENOVKWKEQBELCSVRILT 975
Db 985 SSKLEADIEHLKNKVSEVERNALLAGNERLMD--DLKNNGEN-----IASL 1030
QY 976 NOEEKRRNADILNEKIEBELGRIBEOHRKELEVAQLEQALRIQDIELKSVESNLQVSS 1035
Db 1031 QTEIEKKRAE---NDDLOSKLSVSWSEYENLLISSQTNKSLDKTNQIKYIEKNVQKLL 1087
QY 1036 HTHENENYLLHENC---MLKKEIAMLKLEIATLKHQYOEKENKYFEDIKILKKNAEIQ 1091
Db 1088 DEKQDNVLEBELTSKYKGLGEENAOIKDELLALRKSKKHQHDLCANFVDDLKEKSDALE 1147
QY 1092 MTLKLEBSLTKRASQYSGOLKVLIAENTMLTSKLKE-----K 1129
Db 1148 QLTNEKNE-LIVLEFSQSNNEALVEERSDLANRLSDMKKSLSDSDNIVSVIRSLVRVN 1206
QY 1130 QDKEILEAEIESHHPRLASAVQDHQIIVTSRKSKQEPAPH-----IAGDACLQKRMNVVSS 1185
Db 1207 DELDTLKDKDQSLSTQYSEVQCDRDLDSLKGCESEFNKAVSLRELCTKSEIDVPVSE 1266
QY 1186 TIYNNEVLHQ-PLSEAQRKS-KSLKINLYAGDALR-----ENTLVS----- 1225
Db 1267 ILDDNFVNAGNFSLSLTLVLSLE---NYL-DAFNQVNFKQMLDNLRLTTTDAEFTKV 1322
QY 1226 -----EHAQRD---QR-ETOCQKAEHMYQNEQ-----DNVNKHTEQOESLDQKLFOLQ 1271
Db 1323 ADLEKLOHEHDDWLQIGDLEKALDKSEKNFLKEAEMTENIHSLEEGKEETKEIAELS 1382
QY 1272 SKNMWLOOLVHAHKKAANKKITIDIHFLERKMOHLLKKNBEIFNYNHLLKNRIYQY 1331
Db 1383 SR-----LEDNQLATNKLKNQDLHLNQETIRLKEDEVLKEKESLIISLEESLSN---QR 1431
QY 1332 EKEKA 1336
Db 1432 QKES 1436

RESULT 12
JC5837
364K Golgi complex-associated protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C:Accession: JC5837
R:Toki, C.; Fujiwara, T.; Sohda, M.; Hong, H.S.; Misumi, Y.; Ikehara, Y.
Cell Struct. Funct. 22, 565-577, 1997
A:Title: Identification and characterization of rat 364-kDa Golgi-associated protein rec
F:49-549, 624-1176, 1238-1707, 1763-3114/Domain: coiled-coil leucine zipper #status predic
A:Accession: JC5837
A:Reference number: JC5837; MUID:98093490; PMID:9431462
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-3187 <TOK>
A:Cross-references: UNIPROT:Q63714; DBJ:D25543; NID:G516825; PIDN:BAA05026.1; PID:G5168
C:Comment: This protein plays a role in the formation and maintenance of the characteris
C:Superfamily: giantin
F:49-549, 624-1176, 1238-1707, 1763-3114/Domain: coiled-coil leucine zipper #status predic
F:3165-3187/Domain: membrane anchor #status predicted <MAD>

Query Match 5.5%; Score 381; DB 2; Length 3187;
Best Local Similarity 19.9%; Pred. No. 1e-07;
Matches 332; Conservative 259; Mismatches 563; Indels 516; Gaps 69;
QY 49 EHRTPLMKAL-QCHOEACANILLDSGAD-----INLVGVGNMALHYAVYSEILSVAKL 102
Db 403 KHTSSQTSQSLPDVYNEGVOAVNEESVLSQKRVLELENEKCALLLSSLELELAEKNEKL 462

QY 103 LSHGAVIEVHNKA-----SLTPLLLSITKRSQIIVEFL---LIKNNANANAVNKYKCT 151
Db 463 CSRTITLQAQRAGBAGMVCVSTAGIALLNRSDSSTESQDVLNETFS---QGHKEL 519
QY 152 ALMLAVCHGSSEIVGMLLQ-----NVDVPAAD-ICGVTAESHAVTCGF 194
Db 520 SVLLVEMKEAQEBEATAFLKSQLOGKRKPEGDYEVLDKREKVEGMESEGLSPSTABD--VLC-- 575
QY 195 HHIHEQIMEYIRKLSKNHNTNP--EGTSAGTPD-----EAPLAERTPDTAESLVEKTP 247
Db 576 -----APDKNSVPAVEGEQAGMRDQHGTEAGPL-----NDTGMELNSPQP 617
QY 248 DE-----AAPLV-----BRTPDATBSLVKTPDEPAASLVGEGSDKIQ 284
Db 618 DGVDKLSAPHVCLCHQGELERLKTQVLELSTLHTAKETHEKNLSEKAK-----ETS 670
QY 285 CLEKATSKPQSABETPREITSPAKETSEKFTWPAKPRPKIAWEKKEDTPREIMSPAK 344
Db 671 SLTQLTK-EFKESABEARSJTUTAVCEE-----RDQLLYRVKELDVLGELRAQVR 718
QY 345 ETSEKFTWAAKGRP-----RKIAWEKKETPVKTGCVAR----- 377
Db 719 ELETSLAEAEKQGLDYESQRAQHNLTLTEQIHSLSIEAKSKDVKIETLQRELDGVQLQPS 778
QY 378 -----VTSNNTKYLEKG-RSKVIACPTKESSTKAS----- 406
Db 779 EOGTQIKSQLOLQTKESVLEGAERMKDISKEMBELSQALSOKELEIAQMDQLLEKKK 838
QY 407 -----ANDQF-----PSESQJE 419
Db 839 DVEITLQOTIOBKQOQVTELSFMTKEMVQLNEEKFSLGVEIKTLKEQISLSRAEGAKRE 898
QY 420 EDEEYSCDS---RSLFESSAKIQVCIPESIVQKWEINREVEEPKPSAF---KPAE 472
Db 899 QVESGSAESSPHGPHSSAEPVC-KEALQOELEWLKRESEQRKRLQAALISKELLQ 957
QY 473 MONSVPNKAPFLKNQET---LRADPM--FPPEKOKOYEENSWSDE-----SLCETVSQ 521
Db 958 KVSLEELAKVREESTKDSLRESEKRELEEDSKNKDDPEKYGTSEWRELEVSLRLTISE 1017
QY 522 KQVCL-----PKATHQ-----KEIDKINGKLRESF----- 546
Db 1018 KEVELEGIRDLKEKAAAEELQALVQMTODLQNKTKQIDLLQEEITENQATIQKFTIG 1077
QY 547 -----NKDGLLKATCGMKVSIPTKA-----LELK--DMOTFKABPPGKPSAFEPAT 590
Db 1078 TMDAGDGD SAVKET---SVSSPPRAGGGEHMKPELEGIVLEKEKTQLOKK---LQELAL 1131
QY 591 EMQKSVPNKALE-----LKNEQTRAD-----EILPSESQKQD----- 623
Db 1132 ISRKAILKKAQEKELKEELKEQKDAYRHLQEQFDGQSKENINRAPLURLQAKESTDQ 1191
QY 624 -----YEENSWDTESL-----CETVSQKDVCLPAAHQKEIDKINGKLEGSPVKGLLKA 673
Db 1192 QLPQTGOQEPHTHSGEGLSEGLETPASESDLHAAQSPHGETATLQATVSVIAQIQDL--- 1248
QY 674 NCGMKVSIPTKALELMDMQTFKABPPEKPSAFEPALIMOKSVPNKALELKNQETURADEI 733
Db 1249 ---KEIEVEKELELKISST-----TSELTKKSEEVLLQEQINEQGLIQLNKAASHKAK 1301
QY 734 LPSESQKQDYBESSW---DSESLCETVSQKDVCLPQKATHQK---EIDKINGKLEESPDNDG 788
Db 1302 AHTESQLQKELESSQLKIADLEHL-KTLQPELETILQKHVGQKEEVS YLVGQGE----- 1354
QY 789 FLKAPCRMKVSIPTKALELMDMQTFKABPPEKPSAFEPALIMQ-----KSVPNKALE 840
Db 1355 -----KEQTLTTVQTEMBERQERLIKALHTQLEWQAKHEERLQKQVEICE 1400
QY 841 LKNQOTLRADOMFPSESQKKVE-----ENSWDSLSLRETVS----- 877
Db 1401 LKQPKLEEESEKAKQQLQKRLQAALISRKALKENKLSQLSQLSARDAVEHLTKSLADV 1460
QY 878 -----QKDVCVPK-ATHQKEMDK-----ISGKLKSDS 902

Db	1461	ESQVSVQNEQDALLGKALLQBERDKLIIVEMDKSLLENQSLGSCBSLKLALUGGLTEDK	1520
Qy	903	TSLSKILDTVHSCER--ARELQDHCQRTGRMEQMKKFCVLKXKL-----SEAKEIKS	955
Db	1521	EKLMELESVR-CSKIAESTWQEKH-----KELQKEYEVLLQSVNVSNEARIQH	1571
Qy	956	QLENQKWEQELCSVRITLNOEKEKRNADILNEKIREELGRIBEOHRKELEVYKQOLEQ	1015
Db	1572	VVESVR---QEKQEVYAKLASAESDKRERKQLQDAEQEEMKMKRPAKSKQ--K	1625
Qy	1016	ALRIQDIELKSVESNLNVSHENYLLHENCMLKKEIAMLKLEIATLKHQYOEKENK	1075
Db	1626	ILEEE-ENDRLABAQVPGGASMEALLSSNASLKEELRITILEYKLSKEFE-----	1679
Qy	1076	YPEDIKILKEKNAEIMQTLK-LK-----EESLTKRAS-----QYSGOLKVLIAENTMLTSKLKE	1128
Db	1680	-----ALMAEKVTLSEETRNKLQVEAQELQASLETTEKSDPKDVIEEVTEAVVGKSQ	1734
Qy	1129	KQDKELABIE-----SHHPRLASAVQDHQIVTSRKSQBPAPHIAGDACLQKRM	1179
Db	1735	EQDSLSENAKLEDAEATLANSAPGVSETTFSSHDDINNYLQDQLKGRITAELEMEKQK	1794
Qy	1180	NVDVSTIYNNVHLQPLSEAKSKSLKI-----NLNYAGDALREN-----TLVSEHA-	1228
Db	1795	DRELSQTLNEK--NALLTQISAKSELKLEEEVAKINMLNQIQEELSRVTKLETA	1852
Qy	1229	ORDQRETQCMKEAE-----HMYQNEQDNNVNHTEQOESLDQKLOLOQSKNMWLOOL	1281
Db	1853	BEKDDLEERLMQALNLSIGNYQDVTDAIKN-EQLESEMQNLKRCVSELEBEKQOL	1911
Qy	1282	VNAHKADNKSITTDIHFLEP-----KMHLLKKEKNEI	1317
Db	1912	VREKTVSESE---IRKEYMEKIQAQKPGSGKIHAKELQELLKKEKQEV	1957
RESULT 13			
T18296			
myosin heavy chain - Entamoeba histolytica			
C;Species: Entamoeba histolytica			
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004			
C;Accession: T18296			
R;Guillen, N.			
submitted to the EMBL Data Library, February 1997			
A;Reference number: Z18865			
A;Accession: T18296			
A;Status: preliminary; translated from GB/EMBL/DBJ			
A;Molecule type: DNA			
A;Residues: 1-2139 <GUI>			
A;Cross-references: UNIPROT:Q07569; EMBL:L03534; NID:g1850912; PID:g1850913; PIDN:AAB480			
C;Genetics:			
A;Gene: mhca			
C;Superfamily: myosin heavy chain; myosin motor domain homology			
F;91-780/Domain: myosin motor domain homology <MMO>			
Query Match 5.5%; Score 377.5; DB 2; Length 2139;			
Best Local Similarity 22.0%; Pred. No. 8.9e-08;			
Matches 278; Conservative 179; Mismatches 469; Indels 335; Gaps 52;			
Qy	207	KLKNNHONTNPEGTSACTPDPAAPLAERTPDPTAESLVEKTPDPAALPVRTPTAESL--	264
Db	967	EITELNSQINTLNATVNDKDTIAEMQESIDEKEDEITKLKGIKLEEEKDDLEQDRAD	1026
Qy	265	VEKTPDPAASLVEGTSKIOCLE-KATSGKFEQSAETP---REITSPAKETSEKTPWA	320
Db	1027	VSATKDDIAKLNKIT--ICEBAKDIAKLEQLEDEENKNDNLQOOLKLGETE	1084
Qy	321	KGRPRKIAWEKEDTPEIMSPAKETSEKTPAAKGRPRKIAWEKETPVKTCVARVTS	380
Db	1085	KSLLAAQVAATKASDERDLSQNLK-NEKLT-----TK	1116
Qy	381	NKTKV---LEKGRSMIACTPTKSTKASANDQFPSPESKQEBDEEYSCDSRLFPSSAK	437

C:Genetics:									
A:Gene: zip									
A:Cross-references: FlyBase:FBgn0005634									
A:introns: 10/3; 90/3; 162/3; 215/1; 255/3; 263/3; 303/3; 334/3; 1248/1; 1349/3; 1526/3;									
C:Superfamily: myosin heavy chain; myosin motor domain homology									
C:Keywords: alternative splicing; ATP; nucleotide binding; P-loop									
F:135-855/Domain: myosin motor domain homology <MMOT>									
F:225-232/Region: nucleotide-binding motif A (P-loop)									
Query Match 5.4%; Score 373.5; DB 2; Length 2057;									
Best Local Similarity 20.2%; Pred. No. 1.2e-07;									
Matches 252; Conservative 239; Mismatches 427; Indels 331; Gaps 53;									
Qy	230	PLAERTPTASLSVEKTPDEAAPLIVERTPDTA-----ESLVEKTPDEAASLVEG	278						
Db	927	PLLEVTKQ-EELKVQK-EDELQVREKLTDTLAKNTQYERYKQVALVEK-----TLAEQ	979						
Qy	279	TSDKIQCLEKATSGKF-----EQSAETPREITSPAKETSEKFTWPAKGRPRKIAWEKKE	333						
Db	980	LOAEIELCAEAESRSLMARKQOELEMMOLETRIBESEE-----RVLAGGEEKK	1031						
Qy	334	PTPREIMPAGKTSKFTWAAGRPKIAWEKTPVTKGCVARTSNTKTVLEKGRSM	393						
Db	1032	-LELNIQLEQLEEBEAAQKLEKVLQDAKIKKYBEDLALTDQNKLLKEK---KL	1087						
Qy	394	IACPTKESSTKASANDQRPFSKOEDEEYSCDSRSLFESSAKIQVCIPESIQKMEI	453						
Db	1088	L-----BERANDL---SQTLAEBEEK-----AKHLAKLKAKHEATITE-LEERLHKD	1130						
Qy	454	NREVEEPPKPSAFKPAIEMQNSVFNKAPFLK---NEQTLRADPMFPPEKQKDYENSW	510						
Db	1131	QOQRESDRS-----KKKITEVADLKEQLNERRVQVDEMQAQLAKRE-----	1173						
Qy	511	DSESLCETVSQKD-VCLPKATHQKEIDKINGKLESPPKDGLLKATCGMKVSIPTKALEL	569						
Db	1174	--EELTQTLRLRDEESATKATQAQRELESQLAETQ-----	1208						
Qy	570	KDMQTFKAPPGKPSAFEPATEMOKSVPNKALELNE-----QWTRADILPESKQ---	621						
Db	1209	EDLEAKA-----ARAKAEKVRDLSEELAKNELLSLDTTAAQOELSKREQELA	1261						
Qy	622	TLKKSLEETVNHEGVLADMRHK-----HSQELNSINDQLE-----NLRKAKTVLE	1307						
Qy	679	VSIPTKALELMDMTFKAPPEPKPSAFEPATEMOKSVPNKALELNEQTLRADEILPSES	738						
Db	1308	KAKGTLAEANADLAT-----EL-RSVNSSRQENDRRRRKQASQIAELQV	1350						
Qy	739	KQKDYESSWDSSLCTVSQKDVCLPKATHQK-----EMDKISGKLEDSTSLSKILDT	798						
Db	1351	KLAIEIRARASELQEKTKU-----QOEANIYNQLEEA-----ELKA	1387						
Qy	799	SIPTKALELMDMTFKAPPEPKPSAFEPATEMOKSVPNKALELNEQTLRADQMPFSESK	858						
Db	1388	SAAVKSASNWSQJTEAQ-----QLLEBET---RQKGLSSK	1421						
Qy	859	QKKVENSWDSSLRETVSQKDVCPKATHQK-----EMDKISGKLEDSTSLSKILDT	911						
Db	1422	LRQIES---EKEALQBLEEDDEA---KRNRYERKLAEVTTOMQEIKKKABEDADLAKELR-	1475						
Qy	912	VHSCERARELQDKHCQRTQKMGKKFCVLKKLSEA-KEIKSQLENQKVKWQELCS	970						
Db	1476	-----EKKRLNKD-----TEALERQVKELIAQNDRLDKSKKIQSELEDATIELEAQ---	1523						
Qy	971	VRILNQEEKERNAD-----ILNEKIREELGRIE-----EQRKLELVKQOLEQAL-R	1018						
Db	1524	-RTKVLELEKQKNFDKILAEAKAISEQIAQERDTAEREAREKETKVLVSRELDFAFDK	1582						
Qy	1019	IQDIE--LKSVENLNQVSHTHENENYLLHCNMLAKEIAMLKLEIATLTKHOYQOEKEN--	1074						
Db	1583	IEDLENKRKRTLQNELDLDLANTQGTADQNVHE---LEKAKGALESQIAELKAQNEELEDL	1639						
Qy	1075	KYFEDIKILKEKN-----AELQMTLKLKEESLTKRASQYSGOLKVLIAE-----NTM	1127						
Db	1640	QLTDEDAKLRLLEVNNQALRSQFERDLLAKEEGABEKEKRGVLKQLRDLETDELDEERKQRTAA	1699						
Qy	1122	LTSKLKEQDKKEILLEAEIESHH-----PRLASAVQDHDQIVTSRK-----S	1162						
Db	1700	VASKKLEGLDKETETTWENHNVKVEDALXKAKLQAOVKDALARDAEAAKAAKEELQALS	1759						
Qy	1163	QEPAFIAGDACLQKNNVDVSST-----IYNNEVLHQPLSEARQKRS-----KSLK	1208						
Db	1760	KEADGKVKALEAEVLQLTEDLASSERARRAAETERDELAEBEIANNANKGSLMTIDEKRRLE	1819						
Qy	1209	INLNVAGDALRENTLVSEHAQRDQRETQCCOMKAEAHMYQNEQDNVNVKH-----TEQOE	1261						
Db	1820	ARIATLLEEELSEOSNEVLDRSRKAQLQLEQTTELANEKSNQKQNGRALLERQNK	1879						
Qy	1262	SLDQKLFOLOS-----KMMVLQOOLVHAHKKANXSK	1293						
Db	1880	ELKAKLAETIAQRTKVKATIATLEAKIAKVEEOLENEGKERLLQOK---ANRKMDKKIK	1936						
Qy	1294	-ITTDIHLERQMQHLLKEKNNEIFNYYNNHLKNRIYQYEKE--KARTE	1339						
Db	1937	ELTMWIEDERRHVDQH--KEQMDKMSKRIKLLKKNLDETEEELQKEKTQ	1983						
Search completed: February 19, 2005, 02:12:17									
Job time : 49.9263 secs									

Search completed: February 19, 2005, 02:12:17
Job time : 49.9263 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	6920	100.0	1341	13	US-10-007-805-565	Sequence 565, App	
2	6920	100.0	1341	14	US-10-076-622-565	Sequence 565, App	
3	6920	100.0	1341	14	US-10-177-293-334	Sequence 334, App	
4	6920	100.0	1341	15	US-10-124-805-565	Sequence 565, App	
5	6920	100.0	1341	15	US-10-058-270A-4	Sequence 4, Appli	
6	6892	99.6	1349	13	US-10-007-805-573	Sequence 573, App	
7	6892	99.6	1349	14	US-10-076-622-573	Sequence 573, App	
8	6892	99.6	1349	14	US-10-124-805-573	Sequence 573, App	
9	5130	74.1	1002	9	US-09-604-287A-475	Sequence 475, App	
10	5130	74.1	1002	9	US-09-834-759-475	Sequence 475, App	
11	5130	74.1	1002	10	US-09-551-621-475	Sequence 475, App	
12	5130	74.1	1002	13	US-10-007-805-475	Sequence 475, App	
13	5130	74.1	1002	14	US-10-076-622-475	Sequence 475, App	

DB I MTKRKTINLNQI

Query Match	100.0%;	Score 6920;	DB 13;	Length 1341;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1341;	Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	MTKRKKTINLNIQDAQKRTALHWACVNGHVEEVTVFLVDRKCOLVDLGDGHRTEPLMKALQC	60	
Db	1	MTKRKKTINLNIQDAQKRTALHWACVNGHVEEVTVFLVDRKCOLVDLGDGHRTEPLMKALQC	60	

```
Qy 61 HOEACANILDSGADINLVVYGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 120
Db 61 HOEACANILDSGADINLVVYGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 120
Qy 121 LLSITKRSQIVFELLIKNANANAVNYKCTALMLAVCHGSSEIVGMLLOQNVDFPAADI 180
Db 121 LLSITKRSQIVFELLIKNANANAVNYKCTALMLAVCHGSSEIVGMLLOQNVDFPAADI 180
Qy 181 CGVTAHYAVTCGPHHIEHQIMEYIRKLSKNHONTNPEGTSAGTPDEAAPLAERTPTDAE 240
Db 181 CGVTAHYAVTCGPHHIEHQIMEYIRKLSKNHONTNPEGTSAGTPDEAAPLAERTPTDAE 240
Qy 241 SLVEKTPDEAAPLVERTPDTAESLVEKTPDEAAPLVERTPDTAESLVEKTPDEAAPLVERTPDTAE 300
Db 241 SLVEKTPDEAAPLVERTPDTAESLVEKTPDEAAPLVERTPDTAESLVEKTPDEAAPLVERTPDTAE 300
Qy 301 TPREITSPAKTSEKFTWPAKGRPRKIAWEKKEDTTPREIMSPAKTSEKFTWPAKGRPRK 360
Db 301 TPREITSPAKTSEKFTWPAKGRPRKIAWEKKEDTTPREIMSPAKTSEKFTWPAKGRPRK 360
Qy 361 IAWEKETPVKTCGVARVTSNKTIVLEKGRSKMIACPTKESSTKASANDORFPSPESKQEE 420
Db 361 IAWEKETPVKTCGVARVTSNKTIVLEKGRSKMIACPTKESSTKASANDORFPSPESKQEE 420
Qy 421 DEEYSCDSRSLFESSAKTQVCIPIESIYQKWEINREVEEPPKPSAFKPAIEMQNSVFNK 480
Db 421 DEEYSCDSRSLFESSAKTQVCIPIESIYQKWEINREVEEPPKPSAFKPAIEMQNSVFNK 480
Qy 481 APELKNQOTLRADPMFPPESSKQDYENSWDSSELSCTVSKQDVCLPKATHQKEIDKING 540
Db 481 APELKNQOTLRADPMFPPESSKQDYENSWDSSELSCTVSKQDVCLPKATHQKEIDKING 540
Qy 541 KLEESPNKDGLLKATCGMKVSIPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKA 600
Db 541 KLEESPNKDGLLKATCGMKVSIPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKA 600
Qy 601 LELKNEQTRADEILPSESQKDYENSWDSSELSCTVSKQDVCLPKATHQKEIDKING 660
Db 601 LELKNEQTRADEILPSESQKDYENSWDSSELSCTVSKQDVCLPKATHQKEIDKING 660
Qy 661 LEGSPVKDGLLKANGCMKVISIPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKAL 720
Db 661 LEGSPVKDGLLKANGCMKVISIPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKAL 720
Qy 721 ELKNEQTLRADEILPSESQKDYENSWDSSELSCTVSKQDVCLPKATHQKEIDKING 780
Db 721 ELKNEQTLRADEILPSESQKDYENSWDSSELSCTVSKQDVCLPKATHQKEIDKING 780
Qy 781 ESSPNDGFLKAPCRMKVSIPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKALE 840
Db 781 ESSPNDGFLKAPCRMKVSIPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKALE 840
Qy 841 LKNEQTLRADQMPFPPESKQKYEENSWDSSELSRETQSKDVCPKATHQKEMDKISGKLE 900
Db 841 LKNEQTLRADQMPFPPESKQKYEENSWDSSELSRETQSKDVCPKATHQKEMDKISGKLE 900
Qy 901 DSTSLSKILDTHVSCBRARELQKHCEQRTGKMEQMKKFCVLKKLSEAKETKSOLENO 960
Db 901 DSTSLSKILDTHVSCBRARELQKHCEQRTGKMEQMKKFCVLKKLSEAKETKSOLENO 960
Qy 961 KVKEQELCSVRLTLNOEKEKRNADIINEKIREELGRIEEQRKELEVKQOLEQALRIQ 1020
Db 961 KVKEQELCSVRLTLNOEKEKRNADIINEKIREELGRIEEQRKELEVKQOLEQALRIQ 1020
Qy 1021 DIEKSVESNLNQVSHTHENENYLLHENCWMLKKEIAMLKLEIATLKHQOEBENKYPEDI 1080
Db 1021 DIEKSVESNLNQVSHTHENENYLLHENCWMLKKEIAMLKLEIATLKHQOEBENKYPEDI 1080
Qy 1081 KILKEKNAELQWTLKLKESLTKRASOYSGQLKVLIAENTMLTSKLKEKODKEILEABIE 1140
Db 1081 KILKEKNAELQWTLKLKESLTKRASOYSGQLKVLIAENTMLTSKLKEKODKEILEABIE 1140

RESULT 2
US-10-076-622-565
; Sequence 565, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C11
; CURRENT APPLICATION NUMBER: US/10/076.622
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 565
; LENGTH: 1341
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-622-565

Query Match 100.0%; Score 6920; DB 14; Length 1341;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTKKKTININIQDAQKRTALHWAQVNGHEEVTVFLVDRKCOLDVLGHEHRTPLMKALQC 60
Db 1 MTKKKTININIQDAQKRTALHWAQVNGHEEVTVFLVDRKCOLDVLGHEHRTPLMKALQC 60
Qy 61 HOEACANILDSGADINLVVYGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 120
Db 61 HOEACANILDSGADINLVVYGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 120
Qy 121 LLSITKRSQIVFELLIKNANANAVNYKCTALMLAVCHGSSEIVGMLLOQNVDFPAADI 180
Db 121 LLSITKRSQIVFELLIKNANANAVNYKCTALMLAVCHGSSEIVGMLLOQNVDFPAADI 180
Qy 181 CGVTAHYAVTCGPHHIEHQIMEYIRKLSKNHONTNPEGTSAGTPDEAAPLAERTPTDAE 240
Db 181 CGVTAHYAVTCGPHHIEHQIMEYIRKLSKNHONTNPEGTSAGTPDEAAPLAERTPTDAE 240
Qy 241 SLVEKTPDEAAPLVERTPDTAESLVEKTPDEAAPLVERTPDTAESLVEKTPDEAAPLVERTPDTAE 300
Db 241 SLVEKTPDEAAPLVERTPDTAESLVEKTPDEAAPLVERTPDTAESLVEKTPDEAAPLVERTPDTAE 300
Qy 301 TPREITSPAKTSEKFTWPAKGRPRKIAWEKKEDTTPREIMSPAKTSEKFTWPAKGRPRK 360
Db 301 TPREITSPAKTSEKFTWPAKGRPRKIAWEKKEDTTPREIMSPAKTSEKFTWPAKGRPRK 360
Qy 361 IAWEKETPVKTCGVARVTSNKTIVLEKGRSKMIACPTKESSTKASANDORFPSPESKQEE 420
Db 361 IAWEKETPVKTCGVARVTSNKTIVLEKGRSKMIACPTKESSTKASANDORFPSPESKQEE 420
Qy 421 DEEYSCDSRSLFESSAKTQVCIPIESIYQKWEINREVEEPPKPSAFKPAIEMQNSVFNK 480
Db 421 DEEYSCDSRSLFESSAKTQVCIPIESIYQKWEINREVEEPPKPSAFKPAIEMQNSVFNK 480
```

```
QY 481 APELNEQTLRADPMPFPESKQDYENSWSSESICETVSQKDVCLPKATHQKEIDKING 540
DB 481 APELNEQTLRADPMPFPESKQDYENSWSSESICETVSQKDVCLPKATHQKEIDKING 540
QY 541 KLEESPNKDGLLKATCGMKVSIPTKALELKMOTFKABPPGKPSAFEPATEMOKSVPNKA 600
DB 541 KLEESPNKDGLLKATCGMKVSIPTKALELKMOTFKABPPGKPSAFEPATEMOKSVPNKA 600
QY 601 LELKNEQTRADEILPSESQKDYENSWSSESICETVSQKDVCLPKAAHQKEIDKINGK 660
DB 601 LELKNEQTRADEILPSESQKDYENSWSSESICETVSQKDVCLPKAAHQKEIDKINGK 660
QY 661 LEGSPVKDGLLKANGCMKVSIPTKALELMDMOTFKABPPKPSAFEPATEMOKSVPNKAL 720
DB 661 LEGSPVKDGLLKANGCMKVSIPTKALELMDMOTFKABPPKPSAFEPATEMOKSVPNKAL 720
QY 721 ELKNEQTLRADEILPSESQKDYENSWSSESICETVSQKDVCLPKATHQKEIDKINGKL 780
DB 721 ELKNEQTLRADEILPSESQKDYENSWSSESICETVSQKDVCLPKATHQKEIDKINGKL 780
QY 781 BESPDNDGFLKAPCMKVSIPTKALELMDMOTFKABPPKPSAFEPATEMOKSVPNKALE 840
DB 781 BESPDNDGFLKAPCMKVSIPTKALELMDMOTFKABPPKPSAFEPATEMOKSVPNKALE 840
QY 841 LKNEQTLRADQMPFPESKQKDYENSWSSESICETVSQKDVCLPKATHQKEIDKINGKLE 900
DB 841 LKNEQTLRADQMPFPESKQKDYENSWSSESICETVSQKDVCLPKATHQKEIDKINGKLE 900
QY 901 DSTLSKILDTVHSCERARELOKQCEORTGMEQMKKFCVLLKKLSEAKEIKSOLBNQ 960
DB 901 DSTLSKILDTVHSCERARELOKQCEORTGMEQMKKFCVLLKKLSEAKEIKSOLBNQ 960
QY 961 KVKWEQELCSVRLTLNQEEKRRNADIINEKIREELGRIEOHRKELEVKKOQLEQALRIQ 1020
DB 961 KVKWEQELCSVRLTLNQEEKRRNADIINEKIREELGRIEOHRKELEVKKOQLEQALRIQ 1020
QY 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKEIATMLKLEIATLKHQYQKENKYFEDI 1080
DB 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKEIATMLKLEIATLKHQYQKENKYFEDI 1080
QY 1081 KILKEKNAELQWLTKLKEESITKTRASQYSGQLKVLIAENTMLTSLKLEKQDKEILEABIE 1140
DB 1081 KILKEKNAELQWLTKLKEESITKTRASQYSGQLKVLIAENTMLTSLKLEKQDKEILEABIE 1140
QY 1141 SHHPRLASAVODHQIIVTSRKSQEPAFHAGDACLQRMNVDSVSTIYNNEVLHQPLSEA 1200
DB 1141 SHHPRLASAVODHQIIVTSRKSQEPAFHAGDACLQRMNVDSVSTIYNNEVLHQPLSEA 1200
QY 1201 QRKSKSLKINLYAGDALRENTLVSEHAQRDQRETQCMKEAEMHYQNEQDNVKNHTBQQ 1260
DB 1201 QRKSKSLKINLYAGDALRENTLVSEHAQRDQRETQCMKEAEMHYQNEQDNVKNHTBQQ 1260
QY 1261 ESLDQKLFQLOSKNNWLOQLVHAUKKADNKSKITIDHFLERKQOHLHLEKKEIEIFY 1320
DB 1261 ESLDQKLFQLOSKNNWLOQLVHAUKKADNKSKITIDHFLERKQOHLHLEKKEIEIFY 1320
QY 1321 NNHLKNRIYQYEKEKAETENS 1341
DB 1321 NNHLKNRIYQYEKEKAETENS 1341
```

RESULT 3

```
US-10-177-293-334
; Sequence 334, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
```

```
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Mexic, Funda
; APPLICANT: Sabin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; FILE REFERENCE: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 334
; LENGTH: 1341
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-334
```

```
Query Match 100.0%; Score 6920; DB 14; Length 1341;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MTKRKKITNLNIQDAQKRTALHWACVNGHEEVTFLVDRKQOLDVLDGEHRTPLMKALQC 60
DB 1 MTKRKKITNLNIQDAQKRTALHWACVNGHEEVTFLVDRKQOLDVLDGEHRTPLMKALQC 60
QY 61 HOECANLIDSGADINLVYGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 120
DB 61 HOECANLIDSGADINLVYGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 120
QY 121 LLSITKRSQEIIVEFLLIKNANANAVNKYKCTALMLAVCHGSGSEIVGMLLQQNVVFAADI 180
DB 121 LLSITKRSQEIIVEFLLIKNANANAVNKYKCTALMLAVCHGSGSEIVGMLLQQNVVFAADI 180
QY 181 CGVTAHYAVTCGFFHHIHEQIMYIRKLSKHQNTNPGTSAGTPDEAPLAERTPDTAE 240
DB 181 CGVTAHYAVTCGFFHHIHEQIMYIRKLSKHQNTNPGTSAGTPDEAPLAERTPDTAE 240
QY 241 SLVSEKTPDEAAPLVERTPDTAESLVEKTPDEAASIVGTSKIOCLEKATSGKPFQSAEE 300
DB 241 SLVSEKTPDEAAPLVERTPDTAESLVEKTPDEAASIVGTSKIOCLEKATSGKPFQSAEE 300
QY 301 TPREITSPAKETSEKFTWPAKGRPRKIAWEKKEEDTPREIMSPAKETSEKFTWAAKGRPRK 360
DB 301 TPREITSPAKETSEKFTWPAKGRPRKIAWEKKEEDTPREIMSPAKETSEKFTWAAKGRPRK 360
QY 361 IAWBKETPVKTCGVARTSVNKTVLEKGRSKMIACPTKESSTKASANDORPPSESKEOE 420
DB 361 IAWBKETPVKTCGVARTSVNKTVLEKGRSKMIACPTKESSTKASANDORPPSESKEOE 420
QY 421 DEEYSCDSRSILFESSAKIQCIPESIVOKVMEINREVEEPPKPSAPKPAIEMQNSVFNK 480
DB 421 DEEYSCDSRSILFESSAKIQCIPESIVOKVMEINREVEEPPKPSAPKPAIEMQNSVFNK 480
```

QY 481 AFELKNEOTLRADPMFPPESSKQDYENSWDSSESLCETVSQKDVCLPKATHQKEIDKING 540
DB 481 AFELKNEOTLRADPMFPPESSKQDYENSWDSSESLCETVSQKDVCLPKATHQKEIDKING 540
QY 541 KLESPNKGGLLKATCGMKVSIPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKA 600
DB 541 KLESPNKGGLLKATCGMKVSIPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKA 600
QY 601 LELKNEOTLRADPMFPPESSKQDYENSWDSSESLCETVSQKDVCLPKAAHQKEIDKING 660
DB 601 LELKNEOTLRADPMFPPESSKQDYENSWDSSESLCETVSQKDVCLPKAAHQKEIDKING 660
QY 661 LEGSPVKGGLLKANGCMKVISPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKA 720
DB 661 LEGSPVKGGLLKANGCMKVISPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKA 720
QY 721 ELKNEOTLRADPMFPPESSKQDYENSWDSSESLCETVSQKDVCLPKATHQKEIDKING 780
DB 721 ELKNEOTLRADPMFPPESSKQDYENSWDSSESLCETVSQKDVCLPKATHQKEIDKING 780
QY 781 EESPNDGFLKAPCRMKVISPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKA 840
DB 781 EESPNDGFLKAPCRMKVISPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKA 840
QY 841 LKNEOTLRADQMPFPPSESKQDYENSWDSSESLCETVSQKDVCPKATHQKEMDKISGLE 900
DB 841 LKNEOTLRADQMPFPPSESKQDYENSWDSSESLCETVSQKDVCPKATHQKEMDKISGLE 900
QY 901 DSTLSKILDTVHSCERARELQDCEORTGRMEQMKKFCVLYKKLSEAKBIKSQLENQ 960
DB 901 DSTLSKILDTVHSCERARELQDCEORTGRMEQMKKFCVLYKKLSEAKBIKSQLENQ 960
QY 961 KVKWQELCSVRLTLNQBEEKERNADILNEKIREELGRIEBOHRKELEVKQOLEALRIQ 1020
DB 961 KVKWQELCSVRLTLNQBEEKERNADILNEKIREELGRIEBOHRKELEVKQOLEALRIQ 1020
QY 1021 DIELKSVESNLQVSHENENYLLHENCMLKKEIATLKLHATLKHQYQEKENYFEDI 1080
DB 1021 DIELKSVESNLQVSHENENYLLHENCMLKKEIATLKLHATLKHQYQEKENYFEDI 1080
QY 1081 KILKERNABLOMTLKLKESSLTKRASQYSGQLKVLIAENTMTLSKLEKQDKEILEABIE 1140
DB 1081 KILKERNABLOMTLKLKESSLTKRASQYSGQLKVLIAENTMTLSKLEKQDKEILEABIE 1140
QY 1141 SHHPRLASVQDHDQIVTSRKSOEPAFHAGDACLQKKNVDVSTIYNNEVLHOPLEA 1200
DB 1141 SHHPRLASVQDHDQIVTSRKSOEPAFHAGDACLQKKNVDVSTIYNNEVLHOPLEA 1200
QY 1201 QRSKSLKINLVAGDALRENTLVSEHAORDQRETQCKMEAHMYQNEQDNVNXHTEBOQ 1260
DB 1201 QRSKSLKINLVAGDALRENTLVSEHAORDQRETQCKMEAHMYQNEQDNVNXHTEBOQ 1260
QY 1261 ESLDQKLFQLSQKNMMLQOOLVHAHKADNKSITIDHIFLERKQOHLLKEKNEIFNY 1320
DB 1261 ESLDQKLFQLSQKNMMLQOOLVHAHKADNKSITIDHIFLERKQOHLLKEKNEIFNY 1320
QY 1321 NNHLKNRIYQYEKEKAETENS 1341
DB 1321 NNHLKNRIYQYEKEKAETENS 1341

RESULT 4
US-10-124-805-565
; Sequence 565, Application US/10124805
; Publication No. US20030166022A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C12
; CURRENT APPLICATION NUMBER: US/10/124,805

; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 565
; LENGTH: 1341
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-124-805-565

Query Match 100.0%; Score 6920; DB 14; Length 1341;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTKRKKTTINLINIQAOKRTALHWACVNGHEEVVTVFLVDRKCOLDVLDSGHRTPMLKALOC 60
DB 1 MTKRKKTTINLINIQAOKRTALHWACVNGHEEVVTVFLVDRKCOLDVLDSGHRTPMLKALOC 60
QY 61 HQBACANILIDSGADINLVDVYGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 120
DB 61 HQBACANILIDSGADINLVDVYGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 120
QY 121 LLSITTKRSEQIVFELIKNANANVNYKCTALMLAVCHGSSEIVGMLLQONVDVFAADI 180
DB 121 LLSITTKRSEQIVFELIKNANANVNYKCTALMLAVCHGSSEIVGMLLQONVDVFAADI 180
QY 181 CGVTAHYAVTCGFHHIHEQIMEYIRKLSKNHONTNPEGTSAGTDPDEAAPLABRTPTAE 240
DB 181 CGVTAHYAVTCGFHHIHEQIMEYIRKLSKNHONTNPEGTSAGTDPDEAAPLABRTPTAE 240
QY 241 SLVEKTPDEAAPLIVERTPDTAEISLVEKTPDEAAPLIVERTPDTAEISLVEKTPDEAAPL 300
DB 241 SLVEKTPDEAAPLIVERTPDTAEISLVEKTPDEAAPLIVERTPDTAEISLVEKTPDEAAPL 300
QY 301 TPREITSPAKETSEKFTWPAKGRPRIAWEKKEDTTPREIMSPAKETSEKFTWPAKGRPRIA 360
DB 301 TPREITSPAKETSEKFTWPAKGRPRIAWEKKEDTTPREIMSPAKETSEKFTWPAKGRPRIA 360
QY 361 IAWKKEKTPVKTGCARVTSNKTAVLEKGRSKMIACPTKESSTKASANDORFFSESQEE 420
DB 361 IAWKKEKTPVKTGCARVTSNKTAVLEKGRSKMIACPTKESSTKASANDORFFSESQEE 420
QY 421 DEEYSCDSRSIPFESSAKIQVCIPESIIYQKWEINREVEEPPKPSAFKPAIEMQNSVPNK 480
DB 421 DEEYSCDSRSIPFESSAKIQVCIPESIIYQKWEINREVEEPPKPSAFKPAIEMQNSVPNK 480
QY 481 AFELKNEOTLRADPMFPPESSKQDYENSWDSSESLCETVSQKDVCLPKATHQKEIDKING 540
DB 481 AFELKNEOTLRADPMFPPESSKQDYENSWDSSESLCETVSQKDVCLPKATHQKEIDKING 540
QY 541 KLESPNKGGLLKATCGMKVSIPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKA 600
DB 541 KLESPNKGGLLKATCGMKVSIPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKA 600
QY 601 LELKNEOTLRADPMFPPESSKQDYENSWDSSESLCETVSQKDVCLPKAAHQKEIDKING 660
DB 601 LELKNEOTLRADPMFPPESSKQDYENSWDSSESLCETVSQKDVCLPKAAHQKEIDKING 660
QY 661 LEGSPVKGGLLKANGCMKVISPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKA 720
DB 661 LEGSPVKGGLLKANGCMKVISPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKA 720
QY 721 ELKNEOTLRADPMFPPESSKQDYENSWDSSESLCETVSQKDVCLPKATHQKEIDKING 780
DB 721 ELKNEOTLRADPMFPPESSKQDYENSWDSSESLCETVSQKDVCLPKATHQKEIDKING 780
QY 781 EESPNDGFLKAPCRMKVISPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKA 840
DB 781 EESPNDGFLKAPCRMKVISPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKA 840
QY 841 LKNEOTLRADQMPFPPSESKQDYENSWDSSESLCETVSQKDVCPKATHQKEMDKISGLE 900
DB 841 LKNEOTLRADQMPFPPSESKQDYENSWDSSESLCETVSQKDVCPKATHQKEMDKISGLE 900

Qy 901 DSTSLSKILDTVHSCERARELOKHCEORTGMEQMKKFCVLLKKLSEAKEIKSLENQ 960
Db 901 DSTSLSKILDTVHSCERARELOKHCEORTGMEQMKKFCVLLKKLSEAKEIKSLENQ 960
Qy 961 KVKWEQELCSVRLTLNQEEKRRNADILNEKIREELGRIEBOHRKELEVKQOLEALRIQ 1020
Db 961 KVKWEQELCSVRLTLNQEEKRRNADILNEKIREELGRIEBOHRKELEVKQOLEALRIQ 1020
Qy 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQSEKENKYPDI 1080
Db 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQSEKENKYPDI 1080
Qy 1081 KILKEKNAELQMTLKLKEESITKQASQYSGQLKVLIAENTWLTSLKLEKQDKEILEABIE 1140
Db 1081 KILKEKNAELQMTLKLKEESITKQASQYSGQLKVLIAENTWLTSLKLEKQDKEILEABIE 1140
Qy 1141 SHHPRLASAVODHQIVTSRSQSPAFHAGDACLQKQNVVDSSTIYNNVHLQPLSEA 1200
Db 1141 SHHPRLASAVODHQIVTSRSQSPAFHAGDACLQKQNVVDSSTIYNNVHLQPLSEA 1200
Qy 1201 QRSKSLKINLYAGDALRENTLVSEHAQRDQRETQCMKEAHEMYQNEQDNVKNHTEQQ 1260
Db 1201 QRSKSLKINLYAGDALRENTLVSEHAQRDQRETQCMKEAHEMYQNEQDNVKNHTEQQ 1260
Qy 1261 ESLQKLPOLQSKNMWLOQQLVHAHKKADNKSKITIDHFLERKQMHLLKEKNEEIPNY 1320
Db 1261 ESLQKLPOLQSKNMWLOQQLVHAHKKADNKSKITIDHFLERKQMHLLKEKNEEIPNY 1320
Qy 1321 NNHLKNRIYQYEKKAETENS 1341
Db 1321 NNHLKNRIYQYEKKAETENS 1341

RESULT 5

US-10-058-270A-4
; Sequence 4, Application US/10058270A
; Publication No. US20040029114A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Afar, Daniel
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Breast Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Breast Cancer
; FILE REFERENCE: 018501-0052100S
; CURRENT APPLICATION NUMBER: US/10/058,270A
; CURRENT FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: US 60/263,965
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: US 60/265,928
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: US 09/829,472
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 60/282,698
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 60/288,590
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,443
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: Patent in ver. 2.1
; SEQ ID NO 4
; LENGTH: 1341
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-058-270A-4

Query Match 100.0%; Score 6920; DB 15; Length 1341;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTKKKKTINLNQDAQKRTALHWCVNGHEBVTFLVDRKCOLQDVLGDEHRTPLMKALQC 60
|||||

Db 1 MTKKKKTINLNQDAQKRTALHWCVNGHEBVTFLVDRKCOLQDVLGDEHRTPLMKALQC 60
Qy 61 HQEACANILIDSGADINLVDVYGNMALHYAVYSEILSVVAKLHSHGAVIEVHNKASLTPL 120
Db 61 HQEACANILIDSGADINLVDVYGNMALHYAVYSEILSVVAKLHSHGAVIEVHNKASLTPL 120
Qy 121 LLSITKSEQIVFELLIKNANANAVNKYKCTALMLAVCHGSSSEIVGMLLQONVDVFAADI 180
Db 121 LLSITKSEQIVFELLIKNANANAVNKYKCTALMLAVCHGSSSEIVGMLLQONVDVFAADI 180
Qy 181 CGVTAEHYAVTCGFFHHIHEQIMEYIRKLSKNHNTNPEGTSAGTPDEAAPLAERTPDTAE 240
Db 181 CGVTAEHYAVTCGFFHHIHEQIMEYIRKLSKNHNTNPEGTSAGTPDEAAPLAERTPDTAE 240
Qy 241 SLVEKTPDEAAPLVERTPDTAESLVEKTPDEAASLVEGTSDKIOCLEKATSGKFEQSABE 300
Db 241 SLVEKTPDEAAPLVERTPDTAESLVEKTPDEAASLVEGTSDKIOCLEKATSGKFEQSABE 300
Qy 301 TPREITSPAKETSEKFTWPAKGRPRKIAWEKKEDTPREIMSPAKETSEKFTWPAKGRPRK 360
Db 301 TPREITSPAKETSEKFTWPAKGRPRKIAWEKKEDTPREIMSPAKETSEKFTWPAKGRPRK 360
Qy 361 IAWKKETPVKTCVARVTSNKTIVLEKGRSKMIACPTKESSTKASANDORPPESKQEE 420
Db 361 IAWKKETPVKTCVARVTSNKTIVLEKGRSKMIACPTKESSTKASANDORPPESKQEE 420
Qy 421 DEBYSCDSRSLFESSAKIQVCIPESIQKWEINREVEEPPKPSAPKPAIEMQNSVFNK 480
Db 421 DEBYSCDSRSLFESSAKIQVCIPESIQKWEINREVEEPPKPSAPKPAIEMQNSVFNK 480
Qy 481 AFEKNEQTLRADPMFPPESKQKDYEENSWSLSELCETVSKQDVCLPKATHQEKIDKING 540
Db 481 AFEKNEQTLRADPMFPPESKQKDYEENSWSLSELCETVSKQDVCLPKATHQEKIDKING 540
Qy 541 KLEESPNKDGILLKATCGMKVSIPTKALELDMQTFKAEPPGPKPSAFEPATEMOKSVPNKA 600
Db 541 KLEESPNKDGILLKATCGMKVSIPTKALELDMQTFKAEPPGPKPSAFEPATEMOKSVPNKA 600
Qy 601 LELKNEQTLRADEILPSESQKDYEENSWSLSELCETVSKQDVCLPKAAHQEKIDKING 660
Db 601 LELKNEQTLRADEILPSESQKDYEENSWSLSELCETVSKQDVCLPKAAHQEKIDKING 660
Qy 661 LEGSPVKDGLLKANGCMKVSIPTKALELDMQTFKAEPPPEKPSAFEPATEMOKSVPNKAL 720
Db 661 LEGSPVKDGLLKANGCMKVSIPTKALELDMQTFKAEPPPEKPSAFEPATEMOKSVPNKAL 720
Qy 721 ELKNEQTLRADEILPSESQKDYEENSWSLSELCETVSKQDVCLPKATHQEKIDKING 780
Db 721 ELKNEQTLRADEILPSESQKDYEENSWSLSELCETVSKQDVCLPKATHQEKIDKING 780
Qy 781 EESPDNDGFLKAPCRMKVSIPTKALELDMQTFKAEPPPEKPSAFEPATEMOKSVPNKALE 840
Db 781 EESPDNDGFLKAPCRMKVSIPTKALELDMQTFKAEPPPEKPSAFEPATEMOKSVPNKALE 840
Qy 841 LKNEQTLRADQMFPSESQKDYEENSWSLSELCETVSKQDVCLPKATHQEKIDKING 900
Db 841 LKNEQTLRADQMFPSESQKDYEENSWSLSELCETVSKQDVCLPKATHQEKIDKING 900
Qy 901 DSTSLSKILDTVHSCERARELOKHCEORTGMEQMKKFCVLLKKLSEAKEIKSLENQ 960
Db 901 DSTSLSKILDTVHSCERARELOKHCEORTGMEQMKKFCVLLKKLSEAKEIKSLENQ 960
Qy 961 KVKWEQELCSVRLTLNQEEKRRNADILNEKIREELGRIEBOHRKELEVKQOLEALRIQ 1020
Db 961 KVKWEQELCSVRLTLNQEEKRRNADILNEKIREELGRIEBOHRKELEVKQOLEALRIQ 1020
Qy 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQSEKENKYPDI 1080
Db 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQSEKENKYPDI 1080
Qy 1081 KILKEKNAELQMTLKLKEESITKQASQYSGQLKVLIAENTWLTSLKLEKQDKEILEABIE 1140
Db 1081 KILKEKNAELQMTLKLKEESITKQASQYSGQLKVLIAENTWLTSLKLEKQDKEILEABIE 1140


```
Qy 1141 SHPLASAVODHDOIIVTSRKSQEPAFHAGDACLQKMNVDVSSTIYNNEVLHOPLSEA 1200
Db 1141 SHPLASAVODHDOIIVTSRKSQEPAFHAGDACLQKMNVDVSSTIYNNEVLHOPLSEA 1200
Qy 1201 QKSKSLKINLYAGDALRENTLVSEHAQRDQRETQCMKEAEHMYQNEQDNVKNKHTQEQ 1260
Db 1201 QKSKSLKINLYAGDALRENTLVSEHAQRDQRETQCMKEAEHMYQNEQDNVKNKHTQEQ 1260
Qy 1261 ESLDQKLFQLOSKNMWLOQOLVHAHKADNKSKITIDHFLERKQOHLLKEKNEEIFNY 1320
Db 1261 ESLDQKLFQLOSKNMWLOQOLVHAHKADNKSKITIDHFLERKQOHLLKEKNEEIFNY 1320
Qy 1321 NNHLKNRIYOYEKEKAETENS 1341
Db 1321 NNHLKNRIYOYEKEKAETENS 1341

RESULT 6
US-10-007-805-573
; Sequence 573, Application US/10007805
; Publication No. US20020150581A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margarita
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C10
; CURRENT APPLICATION NUMBER: US/10/007,805
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 593
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 573
; LENGTH: 1349
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-805-573

Query Match 99.6%; Score 6892; DB 13; Length 1349;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1336; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TTKKKTININIQDAKRTALHWACVNGHEEVTVFLVDRKCOLVDLDGHRTPLMKALOCH 61
Db 10 TTKKKTININIQDAKRTALHWACVNGHEEVTVFLVDRKCOLVDLDGHRTPLMKALOCH 69
Qy 62 QACANILDSGADINLVDVGNMALHYAVYSEILSVAKLISHGAVIEVHNKASLTPLL 121
Db 70 QACANILDSGADINLVDVGNMALHYAVYSEILSVAKLISHGAVIEVHNKASLTPLL 129
Qy 122 LSITRSEQIVFLLIKNANANVNYKCTALMLAVCHGSSIEIVGMLOQNVDFPAADIC 181
Db 130 LSITRSEQIVFLLIKNANANVNYKCTALMLAVCHGSLSEIVGMLOQNVDFPAADIC 189
Qy 182 GYTAEHYAVTCGFHHIHEQIMEYIRKLSKNHONTPEGTSAGTDPDEAAPLAERTPTDAES 241
Db 190 GYTAEHYAVTCGFHHIHEQIMEYIRKLSKNHONTPEGTSAGTDPDEAAPLAERTPTDAES 249
Qy 242 LVEKTPDEAAPLVERTPDPTAESLVEKTPDEAASLVEGTSKDIQCLEKATSGKFEQSAEET 301
Db 250 LVEKTPDEAAPLVERTPDPTAESLVEKTPDEAASLVEGTSKDIQCLEKATSGKFEQSAEET 309
Qy 302 PREITSPAKETSEKFTWPAKGRPRKIATWEKEDTPREIMSPAKETSEKFTWAAKGRPRKI 361
```

RESULT 7
US-10-076-622-573

```
; Sequence 573, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C11
; CURRENT APPLICATION NUMBER: US/10/076,622
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 573
; LENGTH: 1349
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-622-573

Query Match      99.6%; Score 6892; DB 14; Length 1349;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1336; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  2  TKRKTINLNIDQAQKRTALHWACVNGHEEVVTFVDRKCCOLDVLDGEHRTPLMKALQCH 61
Db  10  TKRKTINLNIDQAQKRTALHWACVNGHEEVVTFVDRKCCOLDVLDGEHRTPLMKALQCH 69
Qy  62  QEACANILIDSGADINLVYGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 121
Db  70  QEACANILIDSGADINLVYGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 129
Qy  122  LSIITKRSEQIVFLLIKNANANVKNYKCTALMLAVCHSGSSEIVGMLLQQVNVFAADIC 181
Db  130  LSIITKRSEQIVFLLIKNANANVKNYKCTALMLAVCHSGSSEIVGMLLQQVNVFAADIC 189
Qy  182  GVTAHVAVTGCFHHIHOIMEYIRKLSKNHNTNPEGTSAGTDEAPLAERTPTDAES 241
Db  190  GVTAHVAVTGCFHHIHOIMEYIRKLSKNHNTNPEGTSAGTDEAPLAERTPTDAES 249
Qy  242  LVEKTPDEAAPLVERTPTDAESLVKPTDEASLVVEGTSKIOCLEKATSGKFPQSAEET 301
Db  250  LVEKTPDEAAPLVERTPTDAESLVKPTDEASLVVEGTSKIOCLEKATSGKFPQSAEET 309
Qy  302  PREITSPAKETSEKFTWPAKGRPRKIAWEKKEDTPREIMSPAKETSEKFTWAAKGRPRKI 361
Db  310  PREITSPAKETSEKFTWPAKGRPRKIAWEKKEDTPREIMSPAKETSEKFTWAAKGRPRKI 369
Qy  362  AWEKKEPTVKTGCVARVTSNKTIVLEKGRSWMIACTPKESSTKASANDQRPSPSSKQEB 421
Db  370  AWEKKEPTVKTGCVARVTSNKTIVLEKGRSWMIACTPKESSTKASANDQRPSPSSKQEB 429
Qy  422  EYSCDSRSLPSSAKIQVCIPESIQYKWEINREVEEPPKPSAFKPAIMONSVPNKA 481
Db  430  EYSCDSRSLPSSAKIQVCIPESIQYKWEINREVEEPPKPSAFKPAIMONSVPNKA 489
Qy  482  FELKNEQTLRADPMPFPSSKQKDYENSWSDESICETVSQKDVCLPKATHQKEIDKINGK 541
Db  490  FELKNEQTLRADPMPFPSSKQKDYENSWSDESICETVSQKDVCLPKATHQKEIDKINGK 549
Qy  542  LEESPNDGGLLKATCGMKVSIPTKALELMDMTFKAEPPKPSAFEPATEMOKSVPNKAL 601
Db  550  LEESPNDGGLLKATCGMKVSIPTKALELMDMTFKAEPPKPSAFEPATEMOKSVPNKAL 609
Qy  602  ELKNEQTLRADEILPSSKQKDYENSWSDESICETVSQKDVCLPKAAHQKEIDKINGKL 661
Db  610  ELKNEQTLRADEILPSSKQKDYENSWSDESICETVSQKDVCLPKAAHQKEIDKINGKL 669
Qy  662  EGSPVNDGGLLKANGCMKVISIPTKALELMDMTFKAEPPKPSAFEPATEMOKSVPNKALE 721
Db  670  EGSPVNDGGLLKANGCMKVISIPTKALELMDMTFKAEPPKPSAFEPATEMOKSVPNKALE 729
Qy  722  LKNEQTLRADBILPSSKQKDYESSWSDESICETVSQKDVCLPKATHQKEIDKINGKLE 781
Db  730  LKNEQTLRADBILPSSKQKDYESSWSDESICETVSQKDVCLPKATHQKEIDKINGKLE 789
Qy  782  ESPNDGGLFKAPCRMKVISIPTKALELMDMTFKAEPPKPSAFPAIEMOKSVPNKALEL 841
Db  790  ESPNDGGLFKAPCRMKVISIPTKALELMDMTFKAEPPKPSAFPAIEMOKSVPNKALEL 849
Qy  842  KNEQTLRADQMPFPSSKQKDYENSWSDESILRETVSQKDVCPKATHQKEMDKISGKLE 901
Db  850  KNEQTLRADQMPFPSSKQKDYENSWSDESILRETVSQKDVCPKATHQKEMDKISGKLE 909
Qy  902  STLSKILDTVHSCERARELQKHCEORTGWEOMKKKFCVLKKLSEAKIKSOLENOK 961
Db  910  STLSKILDTVHSCERARELQKHCEORTGWEOMKKKFCVLKKLSEAKIKSOLENOK 969
Qy  962  VKWEQELCSVRLTLNQBEERKRNADILNEKIRELGRIEEQHREKELEVKQOLEQALRIQD 1021
Db  970  VKWEQELCSVRLTLNQBEERKRNADILNEKIRELGRIEEQHREKELEVKQOLEQALRIQD 1029
Qy  1022  IELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIK 1081
Db  1030  IELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIK 1089
Qy  1082  ILKEKNAELQWTLKKEESLTKRASQYSGOLKVLIAENTMLTSLKKEKQKEILEAEIES 1141
Db  1090  ILKEKNAELQWTLKKEESLTKRASQYSGOLKVLIAENTMLTSLKKEKQKEILEAEIES 1149
Qy  1142  HHPLASAVODHDOI VTSRKSQEPAFHIIAGDACLQRKMNVDVSTIYNNVHLHOPLSBAQ 1201
Db  1150  HHPLASAVODHDOI VTSRKSQEPAFHIIAGDACLQRKMNVDVSTIYNNVHLHOPLSBAQ 1209
Qy  1202  RKSLSKLINLYAGDALRENTLVSEHAQRDQRETQCMKEAHMYQNEQDNVNNKHTEQOE 1261
Db  1210  RKSLSKLINLYAGDALRENTLVSEHAQRDQRETQCMKEAHMYQNEQDNVNNKHTEQOE 1269
Qy  1262  SLDOQLFQOLSKNMWLOQLVHAHKADNKSKITIDIFLERKMOHLLKKEKBEIFNYN 1321
Db  1270  SLDOQLFQOLSKNMWLOQLVHAHKADNKSKITIDIFLERKMOHLLKKEKBEIFNYN 1329
Qy  1322  NHLKNRIYQYEKEKAETE 1339
Db  1330  NHLKNRIYQYEKEKAETE 1347

RESULT 8
US-10-124-805-573
; Sequence 573, Application US/10124805
; Publication No. US20030166022A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C12
; CURRENT APPLICATION NUMBER: US/10/124,805
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 573
; LENGTH: 1349
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-124-805-573

Query Match      99.6%; Score 6892; DB 14; Length 1349;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1336; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  2  TKRKTINLNIDQAQKRTALHWACVNGHEEVVTFVDRKCCOLDVLDGEHRTPLMKALQCH 61
Db  10  TKRKTINLNIDQAQKRTALHWACVNGHEEVVTFVDRKCCOLDVLDGEHRTPLMKALQCH 69
Qy  62  QEACANILIDSGADINLVYGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 121
```

Db 70 QECANILLISGADINLVDTVGNMALHAYVSEILSVAKLLSHGAVIEVHINKASLTPL 129
Qy 122 LSITKRSQIIVFLLIKNANANAVNKYKCTALMLAVCHGSSEIIVGMLLQOVVDVFAADIC 181
Db 130 LSITKRSQIIVFLLIKNANANAVNKYKCTALMLAVCHGLSEIIVGMLLQOVVDVFAADIC 189
Qy 182 GVTAEHYAVTCGFPHIHQIIMEYIRKLSKNQNTNPEGTSAGTDEAAPLAERTPDTAES 241
Db 190 GVTAEHYAVTCGFPHIHQIIMEYIRKLSKNQNTNPEGTSAGTDEAAPLAERTPDTAES 249
Qy 242 LVEKTPDEAAPIVETPDTAESLVEKTPDEAAPIVETPDTAESLVEKTPDEAAPIVETPDTAES 301
Db 250 LVEKTPDEAAPIVETPDTAESLVEKTPDEAAPIVETPDTAESLVEKTPDEAAPIVETPDTAES 309
Qy 302 PREITSPAKETSEKFTWPAKGRPKIAWEKKEDTPREIMSPAKETSEKFTWPAKGRPKI 361
Db 310 PREITSPAKETSEKFTWPAKGRPKIAWEKKEDTPREIMSPAKETSEKFTWPAKGRPKI 369
Qy 362 AWEKKEPTVTCGVARVTSNKTUVEKGRSKMIACPTKESSTKASANDORPPSSKQBED 421
Db 370 AWEKKEPTVTCGVARVTSNKTUVEKGRSKMIACPTKESSTKASANDORPPSSKQBED 429
Qy 422 BEYSCDSRSLPSSAKIOVCIPESIIYQKWEINREVEPPKPSAFKPAIEMQNSVPNKA 481
Db 430 BEYSCDSRSLPSSAKIOVCIPESIIYQKWEINREVEPPKPSAFKPAIEMQNSVPNKA 489
Qy 482 FELKNEQTLRADMPMPPESSKQKDYENSWSDESLSLCTVSKQDVCLPKATHQKEIDKINGK 541
Db 490 PELKNEQTLRADMPMPPESSKQKDYENSWSDESLSLCTVSKQDVCLPKATHQKEIDKINGK 549
Qy 542 LEESPNDGGLLKATCGMKVSIPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKAL 601
Db 550 LEESPNDGGLLKATCGMKVSIPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKAL 609
Qy 602 ELKNEQTLRADMPMPPESSKQKDYENSWSDESLSLCTVSKQDVCLPKATHQKEIDKINGK 661
Db 610 ELKNEQTLRADMPMPPESSKQKDYENSWSDESLSLCTVSKQDVCLPKATHQKEIDKINGK 669
Qy 662 EGSPVKDGLLKANCQMKVSIPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKAL 721
Db 670 EGSPVKDGLLKANCQMKVSIPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKAL 729
Qy 722 LKNEQTLRADELTPSSKQKDYENSWSDESLSLCTVSKQDVCLPKATHQKEIDKINGKLE 781
Db 730 LKNEQTLRADELTPSSKQKDYENSWSDESLSLCTVSKQDVCLPKATHQKEIDKINGKLE 789
Qy 782 ESPDNDGFLKAPCRMKVSIPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKAL 841
Db 790 ESPDNDGFLKAPCRMKVSIPTKALELMDQMTFKAEPPEKPSAFEPATEMOKSVPNKAL 849
Qy 842 KNEQTLRADMPMPPESSKQKDYENSWSDESLSLCTVSKQDVCLPKATHQKEIDKINGKLE 901
Db 850 KNEQTLRADMPMPPESSKQKDYENSWSDESLSLCTVSKQDVCLPKATHQKEIDKINGKLE 909
Qy 902 STSLSKILDTVHSCERARELQDHCORTGKWEQMKKFCVLKXKLSAKEIKSOLENOK 961
Db 910 STSLSKILDTVHSCERARELQDHCORTGKWEQMKKFCVLKXKLSAKEIKSOLENOK 969
Qy 962 VKWEQELCSVRLTNOQEEKRRNADILNEKIREELGRIBEOHRKLELVQKQLEQALRIOD 1021
Db 970 VKWEQELCSVRLTNOQEEKRRNADILNEKIREELGRIBEOHRKLELVQKQLEQALRIOD 1029
Qy 1022 IELKSVESNLNOVSHTHENYLLHNCMLKKEIAMLKLEIATLKHQYOEKENKYFEDIK 1081
Db 1030 IELKSVESNLNOVSHTHENYLLHNCMLKKEIAMLKLEIATLKHQYOEKENKYFEDIK 1089
Qy 1082 ILKEKNAELQMTLKLKEESLTKRASQYSGOLKVLIAENTMTLASKLEKQDKLEILAEIES 1141
Db 1090 ILKEKNAELQMTLKLKEESLTKRASQYSGOLKVLIAENTMTLASKLEKQDKLEILAEIES 1149
Qy 1142 HHPRLASAVQDHQIIVTSRKSQEPAFHIAAGDACLQRKMNVDVSSITTYNNEVLHQPULSEAQ 1201

Db 1150 HHPRLASAVQDHQIIVTSRKSQEPAFHIAAGDACLQRKMNVDVSSITTYNNEVLHQPULSEAQ 1209
Qy 1202 RKSKSLKINLNYAGDALRENTLVSEHAQRDQRTQOCQKAEHMYQNEQDNVKNKHTEQOE 1261
Db 1210 RKSKSLKINLNYAGDALRENTLVSEHAQRDQRTQOCQKAEHMYQNEQDNVKNKHTEQOE 1269
Qy 1262 SLQDKLQFQSKQNMWLOQQLVHAHKKADNKSKITIDHFLERKQVQHLLKKEKNEEIPFNYN 1321
Db 1270 SLQDKLQFQSKQNMWLOQQLVHAHKKADNKSKITIDHFLERKQVQHLLKKEKNEEIPFNYN 1329
Qy 1322 NHLKNRIYQYEKSKAETE 1339
Db 1330 NHLKNRIYQYEKSKAETE 1347
RESULT 9
US-09-604-287A-475
; Sequence 475, Application US/09604287A
; Patent No. US20020064872A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Heppler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604, 287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
; US-09-604-287A-475
Query Match 74.1%; Score 5130; DB 9; Length 1002;
Best Local Similarity 99.4%; Pred. No. 1.6e-257;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Qy 340 MSPAKETSEKFTWAAKGRPKIAWEKKETPVKTGCVARVTSNKTUVEKGRSKMIACPTK 399
Db 1 MSPAKETSEKFTWAAKGRPKIAWEKKETPVKTGCVARVTSNKTUVEKGRSKMIACPTK 60
Qy 400 ESSTKASANDORPPSSKQKDEEYSCDSRSLPSSAKIOVCIPESIIYQKWEINREVEE 459
Db 61 ESSTKASANDORPPSSKQKDEEYSCDSRSLPSSAKIOVCIPESIIYQKWEINREVEE 120
Qy 460 PPKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPMPPESSKQKDYENSWSDESLSLCTV 519
Db 121 PPKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPMPPESSKQKDYENSWSDESLSLCTV 180
Qy 520 SKQDVCLPKATHQKEIDKINGKLEESPNKQGLLKATCGMKVSIPTKALELMDQMTFKAEP 579
Db 181 SKQDVCLPKATHQKEIDKINGKLEESPNKQGLLKATCGMKVSIPTKALELMDQMTFKAEP 240
Qy 580 PKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPMPPESSKQKDYENSWSDESLSLCTV 639
Db 241 PKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPMPPESSKQKDYENSWSDESLSLCTV 300
Qy 640 QKDVCLPKATHQKEIDKINGKLEESPNKQGLLKATCGMKVSIPTKALELMDQMTFKAEP 699
Db 301 QKDVCLPKATHQKEIDKINGKLEESPNKQGLLKATCGMKVSIPTKALELMDQMTFKAEP 360
Qy 700 EKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPMPPESSKQKDYENSWSDESLSLCTV 759

Db 361 EKPSAFPAIEMQKSVPNKALELKNQETLRADIELPSESKQDYESSWDSLSLCTVSQ 420
Qy 760 KDVCPLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQTFKAEPPE 819
Db 421 KDVCPLPKAXHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQTFKAEPPE 480
Qy 820 KPSAFPAIEMQKSVPNKALELKNQETLRADOMFPSESQKQVSENSWDSLSLCTVSQ 879
Db 481 KPSAFPAIEMQKSVPNKALELKNQETLRADOMFPSESQKQVSENSWDSLSLCTVSQ 540
Qy 880 DVCVPKATHQKEMDKISGKLESDSTLSKILDTVHSCERARELOKDHCEORTGKQEMOMKK 939
Db 541 DVCVPKATHQKEMDKISGKLESDSTLSKILDTVHSCERARELOKDHCEORTGKQEMOMKK 600
Qy 940 FCVLKKLSEAKEIKSQLENQKVWEQELCSVRLTLNQEEKRRNADILNEKIREELGRI 999
Db 601 FCVLKKLSEAKEIKSQLENQKVWEQELCSVRLTLNQEEKRRNADILNEKIREELGRI 660
Qy 1000 BEQHRKELEVQOOLEQALRIODIELKSVESENQVSHTHENENYLLHENCMLKKEIAMLK 1059
Db 661 BEQHRKELEVQOOLEQALRIODIELKSVESENQVSHTHENENYLLHENCMLKKEIAMLK 720
Qy 1060 LEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKLKEESI-TKRASOYSGOLKVLIAEN 1119
Db 721 LEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKLKEESLTKRASQTSQOLKVLIAEN 780
Qy 1120 TMLTSKLKEKQDKEILEAIESHHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQRM 1179
Db 781 TMLTSKLKEKQDKEILEAIESHHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQRM 840
Qy 1180 NVDVSSITYNNEVLHQPSEARQSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 1239
Db 841 NVDVSSITYNNEVLHQPSEARQSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 900
Qy 1240 KEAHEMYQNEODNVNKHTEQOESLDQKLFQOSKNMWLQOQLVHAHKADNKSKITIDIH 1299
Db 901 KEAHEMYQNEODNVNKHTEQOESLDQKLFQOSKNMWLQOQLVHAHKADNKSKITIDIH 960
Qy 1300 FLERKQVHLLKKEKNEEIFNYYNNHLKNRIYQYEKEKAETENS 1341
Db 961 FLERKQVHLLKKEKNEEIFNYYNNHLKNRIYQYEKEKAETENS 1002

RESULT 10

us-09-834-759-475
; Sequence 475, Application US/09834759
; Publication No. US20020085998A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William F.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-834-759-475

Query Match

74.1%; Score 5130; DB 9; Length 1002;

RESULT 11

Best Local Similarity 99.4%; Pred. No. 1.6e-257;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Qy 340 MSPAKETSEKETWAAKGRPRIAWEKKEETPVKTGCVARVTGNKTKVLEKGRSKMIACPTK 399
Db 1 MSPAKETSEKETWAAKGRPRIAWEKKEETPVKTGCVARVTGNKTKVLEKGRSKMIACPTK 60
Qy 400 ESSTKASANDORPPSEKQEEDEYSCDSRSLPESAKIQVCIPESIIYQKVMENREVEE 459
Db 61 ESSTKASANDORPPSEKQEEDEYSCDSRSLPESAKIQVCIPESIIYQKVMENREVEE 120
Qy 460 PPKPSAFKPAIEMQNSVPNKAPFELKNEOTLRADMPPEPKQDYSENSWDSLSLCTV 519
Db 121 PPKPSAFKPAIEMQNSVPNKAPFELKNEOTLRADMPPEPKQDYSENSWDSLSLCTV 180
Qy 520 SQDVCPLPKATHQKEIDKINGKLEESPNKOGLLKATCGMKVSIPTKALELMDQTFKAEP 579
Db 181 SQDVCPLPKATHQKEIDKINGKLEESPNKOGLLKATCGMKVSIPTKALELMDQTFKAEP 240
Qy 580 PGKPSAFEPATEMOKSVPNKALELKNQETWRADIELPSESKQDYSENSWDSLSLCTVS 639
Db 241 PGKPSAFEPATEMOKSVPNKALELKNQETWRADIELPSESKQDYSENSWDSLSLCTVS 300
Qy 640 QKDVCPLPKAXHQKEIDKINGKLEESPVKDGILLKANCGRKVSIPTKALELMDQTFKAEP 699
Db 301 QKDVCPLPKAXHQKEIDKINGKLEESPVKDGILLKANCGRKVSIPTKALELMDQTFKAEP 360
Qy 700 EKPSAFPAIEMQKSVPNKALELKNQETLRADIELPSESKQDYSENSWDSLSLCTVSQ 759
Db 361 EKPSAFPAIEMQKSVPNKALELKNQETLRADIELPSESKQDYSENSWDSLSLCTVSQ 420
Qy 760 KDVCPLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQTFKAEPPE 819
Db 421 KDVCPLPKAXHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQTFKAEPPE 480
Qy 820 KPSAFPAIEMQKSVPNKALELKNQETLRADOMFPSESQKQVSENSWDSLSLCTVSQ 879
Db 481 KPSAFPAIEMQKSVPNKALELKNQETLRADOMFPSESQKQVSENSWDSLSLCTVSQ 540
Qy 880 DVCVPKATHQKEMDKISGKLESDSTLSKILDTVHSCERARELOKDHCEORTGKQEMOMKK 939
Db 541 DVCVPKATHQKEMDKISGKLESDSTLSKILDTVHSCERARELOKDHCEORTGKQEMOMKK 600
Qy 940 FCVLKKLSEAKEIKSQLENQKVWEQELCSVRLTLNQEEKRRNADILNEKIREELGRI 999
Db 601 FCVLKKLSEAKEIKSQLENQKVWEQELCSVRLTLNQEEKRRNADILNEKIREELGRI 660
Qy 1000 BEQHRKELEVQOOLEQALRIODIELKSVESENQVSHTHENENYLLHENCMLKKEIAMLK 1059
Db 661 BEQHRKELEVQOOLEQALRIODIELKSVESENQVSHTHENENYLLHENCMLKKEIAMLK 720
Qy 1060 LEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKLKEESLTKRASOYSGOLKVLIAEN 1119
Db 721 LEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKLKEESLTKRASOYSGOLKVLIAEN 780
Qy 1120 TMLTSKLKEKQDKEILEAIESHHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQRM 1179
Db 781 TMLTSKLKEKQDKEILEAIESHHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQRM 840
Qy 1180 NVDVSSITYNNEVLHQPSEARQSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 1239
Db 841 NVDVSSITYNNEVLHQPSEARQSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 900
Qy 1240 KEAHEMYQNEODNVNKHTEQOESLDQKLFQOSKNMWLQOQLVHAHKADNKSKITIDIH 1299
Db 901 KEAHEMYQNEODNVNKHTEQOESLDQKLFQOSKNMWLQOQLVHAHKADNKSKITIDIH 960
Qy 1300 FLERKQVHLLKKEKNEEIFNYYNNHLKNRIYQYEKEKAETENS 1341
Db 961 FLERKQVHLLKKEKNEEIFNYYNNHLKNRIYQYEKEKAETENS 1002

```

US-09-551-621-475
; Sequence 475, Application US/09551621
; Publication No. US20030104366A1
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 21021.470C5
; CURRENT APPLICATION NUMBER: US/09/551.621
; CURRENT FILING DATE: 2000-04-17
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
; US-09-551-621-475

```

Query Match	74.1%	Score 5130;	DB 10;	Length 1002;
Best Local Similarity	99.4%;	Prod. No. 1.6e-257;		
Matches	996;	Conservative 2;	Mismatches 47;	Indels 0; Gaps 0;
Qy	340	MSPAKETSEKFTWAAGPRPKIAEKKETPVKTCGVARTSNKTKVLEKGRSKMIACPTK	399	
Db	1	MSPAKETSEKFTWAAGPRPKIAEKKETPVKTCGVARTSNKTKVLEKGRSKMIACPTK	60	
Qy	400	ESSTKASANDORFPSESQOEDEEYSCDSRSLFESSAKIQVCIPESIIYQKWEINREVEE	459	
Db	61	ESSTKASANDORFPSESQOEDEEYSCDSRSLFESSAKIQVCIPESIIYQKWEINREVEE	120	
Qy	460	PPKPSAFKPAIEMWNSVPNKAFELKNEOTLRADPMFPPEKQDYENSDWSESLCETV	519	
Db	121	PPKPSAFKPAIEMWNSVPNKAFELKNEOTLRADPMFPPEKQDYENSDWSESLCETV	180	
Qy	520	SQKDVCPLPKATHQKEIDKINGKLEESPNGDGLLKATCGMKVSIPTKALELMDMOTFKAEP	579	
Db	181	SQKDVCPLPKATHQKEIDKINGKLEESPNGDGLLKATCGMKVSIPTKALELMDMOTFKAEP	240	
Qy	580	PGKPSAFPEATEMOKSVPNKALELKNEOTLRADETILPSESQKQDYEESSWDTESICETVS	639	
Db	241	PGKPSAFPEATEMOKSVPNKALELKNEOTLRADETILPSESQKQDYEESSWDTESICETVS	300	
Qy	640	QKDVCPLPKAAHQKEIDKINGKLEGGSPVKDGLLKANGCMKVSIPTKALELMDMOTFKAEP	699	
Db	301	QKDVCPLPKAAHQKEIDKINGKLEGGSPVKDGLLKANGCMKVSIPTKALELMDMOTFKAEP	360	
Qy	700	EKPSAFPEAIEMOKSVPNKALELKNEOTLRADETILPSESQKQDYEESSWDSESLCETVSQ	759	
Db	361	EKPSAFPEAIEMOKSVPNKALELKNEOTLRADETILPSESQKQDYEESSWDSESLCETVSQ	420	
Qy	760	KDVCPLPKATHQKEIDKINGKLEESPNDNGFLKAPCRMVSIPTKALELMDMOTFKAEPEE	819	
Db	421	KDVCPLPKATHQKEIDKINGKLEESPNDNGFLKAPCRMVSIPTKALELMDMOTFKAEPEE	480	
Qy	820	KPSAFPEAIEMOKSVPNKALELKNEOTLRADOMFPPESEKQKVEENSDWSESLRETVSOK	879	
Db	481	KPSAFPEAIEMOKSVPNKALELKNEOTLRADOMFPPESEKQKVEENSDWSESLRETVSOK	540	
Qy	880	DVCVPKATHQKEMDKISGKLEDSLSKILDTVHSCERARELQKQCEORTGKMEQMKK	939	
Db	541	DVCVPKATHQKEMDKISGKLEDSLSKILDTVHSCERARELQKQCEORTGKMEQMKK	600	
Qy	940	FCVLKKKLGSEAKIEIKSOLENQKVKEQEELCSVRLTLNQEEKKRNADIINEKIREELGRI	999	
Db	601	FCVLKKKLGSEAKIEIKSOLENQKVKEQEELCSVRLTLNQEEKKRNADIINEKIREELGRI	660	

```
Db 121 PPKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPESSKQKQYEESSWDSSELCTV 180
Qy 520 SQKDVCLPKATHQKEIDKINGKLEESPNKDGGLKATCGMKVSIPTKALELMDMOTFKABP 579
Db 181 SQKDVCLPKATHQKEIDKINGKLEESPNKDGGLKATCGMKVSIPTKALELMDMOTFKABP 240
Qy 580 PKPSAFEPATEMOKSVPNKALELKNQETLRADQETLPSSEKQKQYEESSWDSSELCTVS 639
Db 241 PKPSAFEPATEMOKSVPNKALELKNQETLRADQETLPSSEKQKQYEESSWDSSELCTVS 300
Qy 640 QKDVCLPKAAHQKEIDKINGKLEESPNKDGGLKANGCMKVSIPTKALELMDMOTFKABP 699
Db 301 QKDVCLPKAAHQKEIDKINGKLEESPNKDGGLKANGCMKVSIPTKALELMDMOTFKABP 360
Qy 700 EKPSAFEPATEMOKSVPNKALELKNQETLRADQETLPSSEKQKQYEESSWDSSELCTVS 759
Db 361 EKPSAFEPATEMOKSVPNKALELKNQETLRADQETLPSSEKQKQYEESSWDSSELCTVS 420
Qy 760 KDVCLPKATHQKEIDKINGKLEESPNKDGGLKAPCRMKVSIPTKALELMDMOTFKABP 819
Db 421 KDVCLPKAXHQKEIDKINGKLEESPNKDGGLKAPCRMKVSIPTKALELMDMOTFKABP 480
Qy 820 KPSAFEPATEMOKSVPNKALELKNQETLRADQETLPSSEKQKQYEESSWDSSELCTVS 879
Db 481 KPSAFEPATEMOKSVPNKALELKNQETLRADQETLPSSEKQKQYEESSWDSSELCTVS 540
Qy 880 QKDVCLPKATHQKEIDKINGKLEESPNKDGGLKANGCMKVSIPTKALELMDMOTFKABP 939
Db 541 QKDVCLPKAXHQKEIDKINGKLEESPNKDGGLKANGCMKVSIPTKALELMDMOTFKABP 600
Qy 940 FCVLKXKLSEAKEIKSOLQENQKQWQELCSVRLTLNQEEKRRNADILNEKIREELGRI 999
Db 601 FCVLKXKLSEAKEIKSOLQENQKQWQELCSVRLTLNQEEKRRNADILNEKIREELGRI 660
Qy 1000 EQHREKLEVKQQLQEQALRIQDIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLK 1059
Db 661 EQHREKLEVKQQLQEQALRIQDIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLK 720
Qy 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSQLKVLIAEN 1119
Db 721 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSQLKVLIAEN 780
Qy 1120 TMLTSKLKEKQDKEILEAEIESHHPRLASAVQDHDQIVTSRKSPAFHAGDACLQKRM 1179
Db 781 TMLTSKLKEKQDKEILEAEIESHHPRLASAVQDHDQIVTSRKSPAFHAGDACLQKRM 840
Qy 1180 NVDVSTIYNNEVLHQPILSEAKRSKSLKINLYAGDALRENTLYSEHAQDQRETQCOM 1239
Db 841 NVDVSTIYNNEVLHQPILSEAKRSKSLKINLYAGDALRENTLYSEHAQDQRETQCOM 900
Qy 1240 KEAEHMYQNEQDNVANKHTEQESLDQKLFLQSKNMWLQOOLVHAHKADNKSKITIDIH 1299
Db 901 KEAEHMYQNEQDNVANKHTEQESLDQKLFLQSKNMWLQOOLVHAHKADNKSKITIDIH 960
Qy 1300 FLERKQVQHLLKKEKNEEIPFNYNHLLKNRIYQYEKEKAETENS 1341
Db 961 FLERKQVQHLLKKEKNEEIPFNYNHLLKNRIYQYEKEKAETENS 1002
```

RESULT 13

```
US-10-076-622-475
; Sequence 475, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; INVENTOR: Sleath, Paul R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C11
; CURRENT APPLICATION NUMBER: US/10/076,622
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 627
```

```
; SOFTWARE: FaSTSeq for Windows Version 4.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 310, 429, 522
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-076-622-475
```

```
Query Match 74.1%; Score 5130; DB 14; Length 1002;
Best Local Similarity 99.4%; Pred. No. 1.6e-257;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
Qy 340 MSPAKETSEKFTTAAAGRPRIAEKKEPTPVKTGCVARVTSNKTAVLEKGRSKMTACPTK 399
Db 1 MSPAKETSEKFTTAAAGRPRIAEKKEPTPVKTGCVARVTSNKTAVLEKGRSKMTACPTK 60
Qy 400 ESSTKASANDQRFPSSEKQDEEYSCDSRSLFESSAKIQVCIPESIIQKVMIEINREVEE 459
Db 61 ESSTKASANDQRFPSSEKQDEEYSCDSRSLFESSAKIQVCIPESIIQKVMIEINREVEE 120
Qy 460 PPKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPESSKQKQYEESSWDSSELCTV 519
Db 121 PPKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPESSKQKQYEESSWDSSELCTV 180
Qy 520 SQKDVCLPKATHQKEIDKINGKLEESPNKDGGLKATCGMKVSIPTKALELMDMOTFKABP 579
Db 181 SQKDVCLPKATHQKEIDKINGKLEESPNKDGGLKATCGMKVSIPTKALELMDMOTFKABP 240
Qy 580 PKPSAFEPATEMOKSVPNKALELKNQETLRADQETLPSSEKQKQYEESSWDSSELCTVS 639
Db 241 PKPSAFEPATEMOKSVPNKALELKNQETLRADQETLPSSEKQKQYEESSWDSSELCTVS 300
Qy 640 QKDVCLPKAAHQKEIDKINGKLEESPNKDGGLKANGCMKVSIPTKALELMDMOTFKABP 699
Db 301 QKDVCLPKAXHQKEIDKINGKLEESPNKDGGLKANGCMKVSIPTKALELMDMOTFKABP 360
Qy 700 EKPSAFEPATEMOKSVPNKALELKNQETLRADQETLPSSEKQKQYEESSWDSSELCTVS 759
Db 361 EKPSAFEPATEMOKSVPNKALELKNQETLRADQETLPSSEKQKQYEESSWDSSELCTVS 420
Qy 760 KDVCLPKATHQKEIDKINGKLEESPNKDGGLKAPCRMKVSIPTKALELMDMOTFKABP 819
Db 421 KDVCLPKAXHQKEIDKINGKLEESPNKDGGLKAPCRMKVSIPTKALELMDMOTFKABP 480
Qy 820 KPSAFEPATEMOKSVPNKALELKNQETLRADQETLPSSEKQKQYEESSWDSSELCTVS 879
Db 481 KPSAFEPATEMOKSVPNKALELKNQETLRADQETLPSSEKQKQYEESSWDSSELCTVS 540
Qy 880 DVCVPKATHQKEMDKISGKLESDSTLSKILDTVHSCERARELQKDHCEQRTGKMEQMKK 939
Db 541 DVCVPKATHQKEMDKISGKLESDSTLSKILDTVHSCERARELQKDHCEQRTGKMEQMKK 600
Qy 940 FCVLKXKLSEAKEIKSOLQENQKQWQELCSVRLTLNQEEKRRNADILNEKIREELGRI 999
Db 601 FCVLKXKLSEAKEIKSOLQENQKQWQELCSVRLTLNQEEKRRNADILNEKIREELGRI 660
Qy 1000 EQHREKLEVKQQLQEQALRIQDIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLK 1059
Db 661 EQHREKLEVKQQLQEQALRIQDIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLK 720
Qy 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSQLKVLIAEN 1119
Db 721 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSQLKVLIAEN 780
Qy 1120 TMLTSKLKEKQDKEILEAEIESHHPRLASAVQDHDQIVTSRKSPAFHAGDACLQKRM 1179
Db 781 TMLTSKLKEKQDKEILEAEIESHHPRLASAVQDHDQIVTSRKSPAFHAGDACLQKRM 840
Qy 1180 NVDVSTIYNNEVLHQPILSEAKRSKSLKINLYAGDALRENTLYSEHAQDQRETQCOM 1239
Db 841 NVDVSTIYNNEVLHQPILSEAKRSKSLKINLYAGDALRENTLYSEHAQDQRETQCOM 900
```


Db 841 NVDVSTIYNNEVLHQLPSEARQSKSLKINLVAGDALRENTLVSEHAQRDQRETQCM 900
Qy 1240 KEAEHMYQNEQNVNKHTEQESLDQKLFQLOSKMMLQOOLVHAHKKADNKSKITIDIH 1299
Db 901 KEAEHMYQNEQNVNKHTEQESLDQKLFQLOSKMMLQOOLVHAHKKADNKSKITIDIH 960
Qy 1300 FLERKQWHLHLLKEKNEEIFNYYNNHLKNRIYQYEKEKAETENS 1341
Db 961 FLERKQWHLHLLKEKNEEIFNYYNNHLKNRIYQYEKEKAETENS 1002

RESULT 14
US-10-124-805-475
; Sequence 475, Application US/10124805
; Publication No. US20030166022A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C12
; CURRENT APPLICATION NUMBER: US/10/124,805
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 310, 429, 522
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-124-805-475

Query Match 74.1%; Score 5130; DB 14; Length 1002;
Best Local Similarity 99.4%; Pred. No. 1.6e-257;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 340 MSPAKETSEKFTWAAKGRPRKIAWEKKEPTVKTGCVARVTSNKTIVLEKGRSKMTACPTK 399
Db 1 MSPAKETSEKFTWAAKGRPRKIAWEKKEPTVKTGCVARVTSNKTIVLEKGRSKMTACPTK 60
Qy 400 ESTKASANDQPPSEKOEDEEYSCDSRSIFESSAKIQVCIPESIQKWEINREVEE 459
Db 61 ESTKASANDQPPSEKOEDEEYSCDSRSIFESSAKIQVCIPESIQKWEINREVEE 120
Qy 460 PPKPSAFPAIEMQNSVFNKAFELKNEQTLRADPMFPPEKQKDYENSWDSLSLCEV 519
Db 121 PPKPSAFPAIEMQNSVFNKAFELKNEQTLRADPMFPPEKQKDYENSWDSLSLCEV 180
Qy 520 SQKDVCLPKATHQKIDKINGKLESPNKGKLLKATCGMKVSIPTKALELMDMTFKAEPP 579
Db 181 SQKDVCLPKATHQKIDKINGKLESPNKGKLLKATCGMKVSIPTKALELMDMTFKAEPP 240
Qy 580 PKPSAFPAIEMQNSVFNKAFELKNEQTLRADPMFPPEKQKDYENSWDSLSLCEV 639
Db 241 PKPSAFPAIEMQNSVFNKAFELKNEQTLRADPMFPPEKQKDYENSWDSLSLCEV 300
Qy 640 QKDVCLPKAAHQKIDKINGKLESPNKGKLLKATCGMKVSIPTKALELMDMTFKAEPP 699
Db 301 QKDVCLPKAAHQKIDKINGKLESPNKGKLLKATCGMKVSIPTKALELMDMTFKAEPP 360
Qy 700 EKPSAFPAIEMQNSVFNKAFELKNEQTLRADPMFPPEKQKDYENSWDSLSLCEV 759
Db 361 EKPSAFPAIEMQNSVFNKAFELKNEQTLRADPMFPPEKQKDYENSWDSLSLCEV 420
Qy 760 KQVCLPKATHQKIDKINGKLESPNKGKLLKATCGMKVSIPTKALELMDMTFKAEPP 819
Db 421 KQVCLPKATHQKIDKINGKLESPNKGKLLKATCGMKVSIPTKALELMDMTFKAEPP 480
Qy 820 KPSAFPAIEMQNSVFNKAFELKNEQTLRADPMFPPEKQKDYENSWDSLSLCEV 879

Db 481 KPSAFPAIEMQNSVFNKAFELKNEQTLRADPMFPPEKQKDYENSWDSLSLCEV 540
Qy 880 DVCVPKATHQKIDKINGKLESPNKGKLLKATCGMKVSIPTKALELMDMTFKAEPP 939
Db 541 DVCVPKATHQKIDKINGKLESPNKGKLLKATCGMKVSIPTKALELMDMTFKAEPP 600
Qy 940 FCVLKKLSEAKETKSOLENQKVKWEQELCSVRLTLNQEEBKRNADILNEKIREELGRI 999
Db 601 FCVLKKLSEAKETKSOLENQKVKWEQELCSVRLTLNQEEBKRNADILNEKIREELGRI 660
Qy 1000 EEQHRKELEVKKQLEQALRIQDIELKSVESNLNQVSHHENTHENYLLHENCWMLKKEIAMLK 1059
Db 661 EEQHRKELEVKKQLEQALRIQDIELKSVESNLNQVSHHENTHENYLLHENCWMLKKEIAMLK 720
Qy 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQWTLKLEBSLTKRASQYSGQLKVLIAEN 1119
Db 721 LEIATLKHQYQEKENKYPEDIKILKEKNAELQWTLKLEBSLTKRASQYSGQLKVLIAEN 780
Qy 1120 TMLTSKLKEXQDKKEILEAEIESHHPRLASAVQDHDQIVTSRKSQEPAFHAGDACLQRM 1179
Db 781 TMLTSKLKEXQDKKEILEAEIESHHPRLASAVQDHDQIVTSRKSQEPAFHAGDACLQRM 840
Qy 1180 NVDVSTIYNNEVLHQLPSEARQSKSLKINLVAGDALRENTLVSEHAQRDQRETQCM 1239
Db 841 NVDVSTIYNNEVLHQLPSEARQSKSLKINLVAGDALRENTLVSEHAQRDQRETQCM 900
Qy 1240 KEAEHMYQNEQNVNKHTEQESLDQKLFQLOSKMMLQOOLVHAHKKADNKSKITIDIH 1299
Db 901 KEAEHMYQNEQNVNKHTEQESLDQKLFQLOSKMMLQOOLVHAHKKADNKSKITIDIH 960
Qy 1300 FLERKQWHLHLLKEKNEEIFNYYNNHLKNRIYQYEKEKAETENS 1341
Db 961 FLERKQWHLHLLKEKNEEIFNYYNNHLKNRIYQYEKEKAETENS 1002

RESULT 15
US-09-834-759-493
; Sequence 493, Application US/09834759
; Publication No. US20020085998A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, JIANGCHUN
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 493
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(1095)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-834-759-493

Query Match 74.1%; Score 5130; DB 9; Length 1095;
Best Local Similarity 99.4%; Pred. No. 1.8e-257;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 340 MSPAKETSEKFTWAAKGRPRKIAWEKKEPTVKTGCVARVTSNKTIVLEKGRSKMTACPTK 399
Db 94 MSPAKETSEKFTWAAKGRPRKIAWEKKEPTVKTGCVARVTSNKTIVLEKGRSKMTACPTK 153

Qy	400	ESSTKASANDORPP	SESKQBEDEBYS	CD	SRSLFP	ESSAKI	OV	CIP	ESIYQVK	MEIN	REVBE	459
Db	154	ESSTKASANDORPP	SESKQBEDEBYS	CD	SRSLFP	ESSAKI	OV	CIP	ESIYQVK	MEIN	REVBE	213
Qy	460	PPKPSAFKPAI	EMONSVP	NKAFEL	KNEQTL	RAD	MPFP	PP	SESKOKDY	EENS	WDSE	519
Db	214	PPKPSAFKPAI	EMONSVP	NKAFEL	KNEQTL	RAD	MPFP	PP	SESKOKDY	EENS	WDSE	273
Qy	520	SQDVCLP	KATHQ	KEIDK	INGKLES	PNKDGL	LKAT	CGMKV	SIP	TKALEL	KDMQ	579
Db	274	SQDVCLP	KATHQ	KEIDK	INGKLES	PNKDGL	LKAT	CGMKV	SIP	TKALEL	KDMQ	333
Qy	580	PKGSAP	EPATEM	QKSV	PNKALEL	KNEQ	TWRAD	BI	LP	SESKOKDY	EENS	639
Db	334	PKGSAP	EPATEM	QKSV	PNKALEL	KNEQ	TWRAD	BI	LP	SESKOKDY	EENS	393
Qy	640	QKDVL	PKAAH	QKEID	KINGK	LEGSPV	KDGL	LK	ANC	GMKV	SIP	699
Db	394	QKDVL	PKAAH	QKEID	KINGK	LEGSPV	KDGL	LK	ANC	GMKV	SIP	453
Qy	700	EKPSAF	PAIEM	QKSV	PNKALEL	KNEQ	TWRAD	BI	LP	SESKOKDY	EENS	759
Db	454	EKPSAF	PAIEM	QKSV	PNKALEL	KNEQ	TWRAD	BI	LP	SESKOKDY	EENS	513
Qy	760	KDVCLP	KATHQ	KEIDK	INGK	LES	PND	GF	LKAP	CRMKV	SIP	819
Db	514	KDVCLP	KATHQ	KEIDK	INGK	LES	PND	GF	LKAP	CRMKV	SIP	573
Qy	820	KPSAF	PAIEM	QKSV	PNKALEL	KNEQ	TWRAD	BI	LP	SESKOKDY	EENS	879
Db	574	KPSAF	PAIEM	QKSV	PNKALEL	KNEQ	TWRAD	BI	LP	SESKOKDY	EENS	633
Qy	880	DVCP	KATHQ	KEIDK	INGK	LES	PND	GF	LKAP	CRMKV	SIP	939
Db	634	DVCP	KATHQ	KEIDK	INGK	LES	PND	GF	LKAP	CRMKV	SIP	693
Qy	940	FCVL	KKKL	SEAK	EIK	SOLEN	QKV	WE	QEL	CS	VR	999
Db	694	FCVL	KKKL	SEAK	EIK	SOLEN	QKV	WE	QEL	CS	VR	753
Qy	1000	BEQHR	KE	LEV	QO	LE	QAL	R	IO	DI	EL	1059
Db	754	BEQHR	KE	LEV	QO	LE	QAL	R	IO	DI	EL	813
Qy	1060	LEI	ATL	KHO	Y	Q	E	K	N	Y	P	1119
Db	814	LEI	ATL	KHO	Y	Q	E	K	N	Y	P	873
Qy	1120	TML	T	S	K	U	K	E	D	Q	P	1179
Db	874	TML	T	S	K	U	K	E	D	Q	P	933
Qy	1180	NVD	V	S	T	I	N	N	E	V	L	1239
Db	934	NVD	V	S	T	I	N	N	E	V	L	993
Qy	1240	KEA	E	H	M	Y	Q	E	O	N	D	1299
Db	994	KEA	E	H	M	Y	Q	E	O	N	D	1053
Qy	1300	F	L	E	R	M	O	H	L	L	K	1341
Db	1054	F	L	E	R	M	O	H	L	L	K	1095

Search completed: February 19, 2005, 02:33:46
Job time : 130.953 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:57:06 ; Search time 36.5216 Seconds
(without alignments)
2740.961 Million cell updates/sec

Title: US-09-602-362E-23

Perfect score: 6920

Sequence: 1 MTRKKTINLIQDAQRKA.....NHLKNRIYQYEKKAETENS 1341

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5130	74.1	1002	4	US-09-620-405B-475
2	5130	74.1	1002	4	US-09-604-287A-475
3	5130	74.1	1002	4	US-09-834-759-475
4	5130	74.1	1002	4	US-09-590-751A-475
5	5130	74.1	1002	4	US-09-551-621-475
6	5130	74.1	1095	4	US-09-620-405B-493
7	5130	74.1	1095	4	US-09-834-759-493
8	3252.5	47.0	650	4	US-09-620-405B-469
9	3252.5	47.0	650	4	US-09-433-826B-469
10	3252.5	47.0	650	4	US-09-604-287A-469
11	3252.5	47.0	650	4	US-09-834-759-469
12	3252.5	47.0	650	4	US-09-590-751A-469
13	3252.5	47.0	650	4	US-09-551-621-469
14	3252.5	47.0	743	4	US-09-620-405B-494
15	3252.5	47.0	743	4	US-09-834-759-494
16	2589	37.4	512	4	US-09-451-739H-16
17	2270	32.8	466	4	US-09-620-405B-472
18	2270	32.8	466	4	US-09-433-826B-472
19	2270	32.8	466	4	US-09-604-287A-472
20	2270	32.8	466	4	US-09-834-759-472
21	2270	32.8	466	4	US-09-590-751A-472
22	2270	32.8	466	4	US-09-551-621-472
23	2256	32.6	445	4	US-09-620-405B-473
24	2256	32.6	445	4	US-09-433-826B-473
25	2256	32.6	445	4	US-09-604-287A-473
26	2256	32.6	445	4	US-09-834-759-473
27	2256	32.6	445	4	US-09-590-751A-473

28	2256	32.6	445	4	US-09-551-621-473	Sequence 473, App
29	2212	32.0	432	4	US-09-389-681-181	Sequence 181, App
30	2212	32.0	432	4	US-09-620-405B-181	Sequence 181, App
31	2212	32.0	432	4	US-09-339-338-181	Sequence 181, App
32	2212	32.0	432	4	US-09-433-826B-181	Sequence 181, App
33	2212	32.0	432	4	US-09-604-287A-181	Sequence 181, App
34	2212	32.0	432	4	US-09-285-480-181	Sequence 181, App
35	2212	32.0	432	4	US-09-834-759-181	Sequence 181, App
36	2212	32.0	432	4	US-09-590-751A-181	Sequence 181, App
37	2212	32.0	432	4	US-09-551-621-181	Sequence 181, App
38	1583.5	22.9	410	4	US-09-620-405B-495	Sequence 495, App
39	1583.5	22.9	410	4	US-09-834-759-495	Sequence 495, App
40	1576	22.8	317	4	US-09-389-681-176	Sequence 176, App
41	1576	22.8	317	4	US-09-620-405B-176	Sequence 176, App
42	1576	22.8	317	4	US-09-339-338-176	Sequence 176, App
43	1576	22.8	317	4	US-09-433-826B-176	Sequence 176, App
44	1576	22.8	317	4	US-09-604-287A-176	Sequence 176, App
45	1576	22.8	317	4	US-09-285-480-176	Sequence 176, App

ALIGNMENTS

RESULT 1

US-09-620-405B-475
; Sequence 475, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620.405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-620-405B-475

Query Match						74.1%; Score 5130; DB 4; Length 1002;
Best Local Similarity						99.4%; Pred. No. 0;
Matches						996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Qy	340	MSPAKETSEKFTWAAGPRKIAWEKETPVKTGCVARVTGNKTKVLEKGRSKMIACPTK	399			
Db	1	MSPAKETSEKFTWAAGPRKIAWEKETPVKTGCVARVTGNKTKVLEKGRSKMIACPTK	60			
Qy	400	ESSTKASANDQRPSPSKQDEEYSCDSRLFPSSAKIQVCIPESIQKVMENREVVEE	459			
Db	61	ESSTKASANDQRPSPSKQDEEYSCDSRLFPSSAKIQVCIPESIQKVMENREVVEE	120			
Qy	460	PPKPSAPKPAIEMQNSVPNKAFELKNEQTIRADPMFPSPSKQDYENSWDSLSLCTV	519			
Db	121	PPKPSAPKPAIEMQNSVPNKAFELKNEQTIRADPMFPSPSKQDYENSWDSLSLCTV	180			
Qy	520	SQKDVCLPKATHQEKIDKINGKLBSFNKDGKLLATCGMKVSIPTKALELXDMQTFKAEP	579			
Db	181	SQKDVCLPKATHQEKIDKINGKLBSFNKDGKLLATCGMKVSIPTKALELXDMQTFKAEP	240			
Qy	580	PGKPSAFEPATEMOKSVENKALELKNQETWRADEILPSPSKQDYENSWDSLSLCTV	639			

QY 460 PPKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPPESKQKDYENSWSLSLCTV 519
DB 121 PPKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPPESKQKDYENSWSLSLCTV 180
QY 520 SOKVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 579
DB 181 SOKVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 240
QY 580 PKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPPESKQKDYENSWSLSLCTV 639
DB 241 PKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPPESKQKDYENSWSLSLCTV 300
QY 640 QKDVCLPKAAHOKKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 699
DB 301 QKDVCLPKAAHOKKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 360
QY 700 EKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPPESKQKDYENSWSLSLCTV 759
DB 361 EKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPPESKQKDYENSWSLSLCTV 420
QY 760 KDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 819
DB 421 KDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 480
QY 820 KPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPPESKQKDYENSWSLSLCTV 879
DB 481 KPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPPESKQKDYENSWSLSLCTV 540
QY 880 DVCVPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 939
DB 541 DVCVPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 600
QY 940 FCVLKXKLSEAKKEIKSOLQENQKWEQELCSVRLTLNQEERKRRNADILNEKIREELGRI 999
DB 601 FCVLKXKLSEAKKEIKSOLQENQKWEQELCSVRLTLNQEERKRRNADILNEKIREELGRI 660
QY 1000 EQHRKELEVKQOLEQALRIODIELKSVESNLNOVSHTHENYLLHENCMLKKEIAMLK 1059
DB 661 EQHRKELEVKQOLEQALRIODIELKSVESNLNOVSHTHENYLLHENCMLKKEIAMLK 720
QY 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSQGLKVLIEN 1119
DB 721 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSQGLKVLIEN 780
QY 1120 TMLTSKLKEKQDKEILEAEIESHHPRLASAVODHDQIVTSRKSQEPAFHAGDACLQRM 1179
DB 781 TMLTSKLKEKQDKEILEAEIESHHPRLASAVODHDQIVTSRKSQEPAFHAGDACLQRM 840
QY 1180 NVDVSTTYNNEVLHQPILSEAKRSKSLKINLYAGDALRENTLVSEHAQRDQRETQCM 1239
DB 841 NVDVSTTYNNEVLHQPILSEAKRSKSLKINLYAGDALRENTLVSEHAQRDQRETQCM 900
QY 1240 KEAEMYQNEQDNVKNKTEQESLQKLFQLOSKNMWLOQOLVHAHKADNKSITIDIH 1299
DB 901 KEAEMYQNEQDNVKNKTEQESLQKLFQLOSKNMWLOQOLVHAHKADNKSITIDIH 960
QY 1300 FLERKQVQHLLKKEKEEIEFNYNHLLKNRIYQYEKEKAETENS 1341
DB 961 FLERKQVQHLLKKEKEEIEFNYNHLLKNRIYQYEKEKAETENS 1002

RESULT 5

US-09-551-621-475
; Sequence 475, Application US/09551621
; Patent No. 6825175
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C5
; CURRENT APPLICATION NUMBER: US/09/551.621
; CURRENT FILING DATE: 2000-04-17
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-551-621-475

Query Match 74.1%; Score 5130; DB 4; Length 1002;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 340 MSPAKETSEKFTWAAKGRPRKIAWEKKEETPVKTGCVARVTSNKTVKLEKGRSKMIACPTK 399
DB 1 MSPAKETSEKFTWAAKGRPRKIAWEKKEETPVKTGCVARVTSNKTVKLEKGRSKMIACPTK 60
QY 400 ESSTKASANDORFPSESKQBEDEYSCDSRLFPSSAKIQVCIPEISYQKWEINREVVEE 459
DB 61 ESSTKASANDORFPSESKQBEDEYSCDSRLFPSSAKIQVCIPEISYQKWEINREVVEE 120
QY 460 PPKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPPESKQKDYENSWSLSLCTV 519
DB 121 PPKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPPESKQKDYENSWSLSLCTV 180
QY 520 SOKDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 579
DB 181 SOKDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 240
QY 580 PKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPPESKQKDYENSWSLSLCTV 639
DB 241 PKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPPESKQKDYENSWSLSLCTV 300
QY 640 QKDVCLPKAAHOKKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 699
DB 301 QKDVCLPKAAHOKKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 360
QY 700 EKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPPESKQKDYENSWSLSLCTV 759
DB 361 EKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPPESKQKDYENSWSLSLCTV 420
QY 760 KDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 819
DB 421 KDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 480
QY 820 KPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPPESKQKDYENSWSLSLCTV 879
DB 481 KPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPPESKQKDYENSWSLSLCTV 540
QY 880 DVCVPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 939
DB 541 DVCVPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 600
QY 940 FCVLKXKLSEAKKEIKSOLQENQKWEQELCSVRLTLNQEERKRRNADILNEKIREELGRI 999
DB 601 FCVLKXKLSEAKKEIKSOLQENQKWEQELCSVRLTLNQEERKRRNADILNEKIREELGRI 660
QY 1000 EQHRKELEVKQOLEQALRIODIELKSVESNLNOVSHTHENYLLHENCMLKKEIAMLK 1059
DB 661 EQHRKELEVKQOLEQALRIODIELKSVESNLNOVSHTHENYLLHENCMLKKEIAMLK 720
QY 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSQGLKVLIEN 1119
DB 721 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSQGLKVLIEN 780
QY 1120 TMLTSKLKEKQDKEILEAEIESHHPRLASAVODHDQIVTSRKSQEPAFHAGDACLQRM 1179

```
Db 781 TMLTSKLEKODKEILEAEIESHPRLASAVQDHDQIVTSRKSQBPAHIAGDACLQRM 840
Qy 1180 NVDVSTIYNNEVLHQPSEAKRSKSLKINLYAGDALRENTLVSEHAQRDQRETQCM 1239
Db 841 NVDVSTIYNNEVLHQPSEAKRSKSLKINLYAGDALRENTLVSEHAQRDQRETQCM 900
Qy 1240 KEAHEMYQNEQDNVKNKTEQESLDQKLFQLOSKNMWLOQQLVHAHKADNKSKITIDH 1299
Db 901 KEAHEMYQNEQDNVKNKTEQESLDQKLFQLOSKNMWLOQQLVHAHKADNKSKITIDH 960
Qy 1300 FLERKQMHLLKKEKNEEIFNYYNNHLKNRIYQYEKEKAETENS 1341
Db 961 FLERKQMHLLKKEKNEEIFNYYNNHLKNRIYQYEKEKAETENS 1002

RESULT 6
US-09-620-405B-493
; Sequence 493, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jianshun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 493
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: variant
; LOCATION: (1)...(1095)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-620-405B-493

Query Match 74.1%; Score 5130; DB 4; Length 1095;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 340 MSPAKETSEKFTWAAKGRPRKIAWEKKEKTPVKTGCVARVTSNKTIVLEKGRSKMIACPTK 399
Db 94 MSPAKETSEKFTWAAKGRPRKIAWEKKEKTPVKTGCVARVTSNKTIVLEKGRSKMIACPTK 153
Qy 400 ESSTKASANDORFPSESQDEEYSCDSRSLFESSAKIOVCIPESYIYOKWEINREYEE 459
Db 154 ESSTKASANDORFPSESQDEEYSCDSRSLFESSAKIOVCIPESYIYOKWEINREYEE 213
Qy 460 PPKPSAFKPAHEMNSVPNKAFELKNEQTLRADPMFPPEKQKDYENSWDSLSLCTV 519
Db 214 PPKPSAFKPAHEMNSVPNKAFELKNEQTLRADPMFPPEKQKDYENSWDSLSLCTV 273
Qy 520 SOKVCLPKATHQKEIDKINGKLESPNKGGLKATCGMKVSIPTKALELMDQMTFKABP 579
Db 274 SOKVCLPKATHQKEIDKINGKLESPNKGGLKATCGMKVSIPTKALELMDQMTFKABP 333
Qy 580 PKPSAFEPATMOKSVPNKALELNEQTLRADEILPSESQKDYENSWDSLSLCTV 639
Db 334 PKPSAFEPATMOKSVPNKALELNEQTLRADEILPSESQKDYENSWDSLSLCTV 393
Qy 640 QKDVCLPKAAHQKEIDKINGKLESPVKGGLKANCCKMKVSIPTKALELMDQMTFKABP 699
Db 394 QKDVCLPKAAHQKEIDKINGKLESPVKGGLKANCCKMKVSIPTKALELMDQMTFKABP 453
Qy 700 EKPSAFEPATMOKSVPNKALELNEQTLRADEILPSESQKDYENSWDSLSLCTV 759
```

```
Db 454 EKPSAFEPATMOKSVPNKALELNEQTLRADEILPSESQKDYENSWDSLSLCTV 513
Qy 760 KDVCLPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDQMTFKABP 819
Db 514 KDVCLPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDQMTFKABP 573
Qy 820 KPSAFEPATMOKSVPNKALELNEQTLRADQMPFSESQKDYENSWDSLSLCTV 879
Db 574 KPSAFEPATMOKSVPNKALELNEQTLRADQMPFSESQKDYENSWDSLSLCTV 633
Qy 880 DVCVPKATHQKEMDKISCKLEDSTSLKILDTVHSCERARELOKDHCEQRTGKMEQMKK 939
Db 634 DVCVPKATHQKEMDKISCKLEDSTSLKILDTVHSCERARELOKDHCEQRTGKMEQMKK 693
Qy 940 FCVLKCKLSEAKETKSOLEKQVWEQELCSVRILTQNEEEKRRNADILNEKIREELGRI 999
Db 694 FCVLKCKLSEAKETKSOLEKQVWEQELCSVRILTQNEEEKRRNADILNEKIREELGRI 753
Qy 1000 EEOHKELEVKQOLEQALRIQDIELKSVESNLQVSHTHENYLLHNCMLKKEIAMLK 1059
Db 754 EEOHKELEVKQOLEQALRIQDIELKSVESNLQVSHTHENYLLHNCMLKKEIAMLK 813
Qy 1060 LEIATLKHQYQEKENKYPEDIKILKEKVAELQMTLKLKEESLTKRASQYSGOLKYLIAEN 1119
Db 814 LEIATLKHQYQEKENKYPEDIKILKEKVAELQMTLKLKEESLTKRASQYSGOLKYLIAEN 873
Qy 1120 TMLTSKLEKODKEILEAEIESHPRLASAVQDHDQIVTSRKSQBPAHIAGDACLQRM 1179
Db 874 TMLTSKLEKODKEILEAEIESHPRLASAVQDHDQIVTSRKSQBPAHIAGDACLQRM 933
Qy 1180 NVDVSTIYNNEVLHQPSEAKRSKSLKINLYAGDALRENTLVSEHAQRDQRETQCM 1239
Db 934 NVDVSTIYNNEVLHQPSEAKRSKSLKINLYAGDALRENTLVSEHAQRDQRETQCM 993
Qy 1240 KEAHEMYQNEQDNVKNKTEQESLDQKLFQLOSKNMWLOQQLVHAHKADNKSKITIDH 1299
Db 994 KEAHEMYQNEQDNVKNKTEQESLDQKLFQLOSKNMWLOQQLVHAHKADNKSKITIDH 1053
Qy 1300 FLERKQMHLLKKEKNEEIFNYYNNHLKNRIYQYEKEKAETENS 1341
Db 1054 FLERKQMHLLKKEKNEEIFNYYNNHLKNRIYQYEKEKAETENS 1095

RESULT 7
US-09-834-759-493
; Sequence 493, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jianshun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 493
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: variant
; LOCATION: (1)...(1095)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-834-759-493
```



```
Db 601 FCVLKXKLESEAKIQLSELENQKVKWEQLCSVRFLTL 637
|||||
RESULT 9
US-09-433-826B-469
; Sequence 469, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C4
; CURRENT APPLICATION NUMBER: US/09/433,826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (310)
; NAME/KEY: Xaa = Any Amino Acid<221> unsure
; LOCATION: (429)
; NAME/KEY: Xaa = Any Amino Acid<221> unsure
; LOCATION: (522)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-433-826B-469

Query Match 47.0%; Score 3252.5; DB 4; Length 650;
Best Local Similarity 98.9%; Pred. No. 7e-215;
Matches 630; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 340 MSPAKETSEKFTWAAGRPRIAKWEKETPVKTGCVARVTSNKTVKLEGRSKMIACPTK 399
Db 1 MSPAKETSEKFTWAAGRPRIAKWEKETPVKTGCVARVTSNKTVKLEGRSKMIACPTK 60
Qy 400 ESSTKASANDORPSESQDEDEYSCDSRLPSSAKIQVICIPESIQKMEINREVEE 459
Db 61 ESSTKASANDORPSESQDEDEYSCDSRLPSSAKIQVICIPESIQKMEINREVEE 120
Qy 460 PPKPSAFKPAIEMQNSVPNKAFELKNEOTLRADMPPEPSKQDYBENSWSLSLCEV 519
Db 121 PPKPSAFKPAIEMQNSVPNKAFELKNEOTLRADMPPEPSKQDYBENSWSLSLCEV 180
Qy 520 SQKDVCLPKATHQKEIDKINGKLESPNKGILLKATCGMKVSIPTKALELMDMTFFKAEP 579
Db 181 SQKDVCLPKATHQKEIDKINGKLESPNKGILLKATCGMKVSIPTKALELMDMTFFKAEP 240
Qy 580 PGKPSAFEPATEMOKSVNKALELKNQETWRADEILPSESQKQDYBENSWSLSLCEV 639
Db 241 PGKPSAFEPATEMOKSVNKALELKNQETWRADEILPSESQKQDYBENSWSLSLCEV 300
Qy 640 QKDVCLPKAAHQKEIDKINGKLESPVKGILLKANCCKVSIPTKALELMDMTFFKAEP 699
Db 301 QKDVCLPKAAHQKEIDKINGKLESPVKGILLKANCCKVSIPTKALELMDMTFFKAEP 360
Qy 700 EKPSAFEPATEMOKSVNKALELKNQETWRADEILPSESQKQDYBENSWSLSLCEV 759
Db 361 EKPSAFEPATEMOKSVNKALELKNQETWRADEILPSESQKQDYBENSWSLSLCEV 420
Qy 760 KDVCLPKATHQKEIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMTFFKAEP 819
Db 421 KDVCLPKATHQKEIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMTFFKAEP 480
Qy 820 KPSAFEPATEMOKSVNKALELKNQETLRADMPPEPSKQDYBENSWSLSLCEV 879
Db 481 KPSAFEPATEMOKSVNKALELKNQETLRADMPPEPSKQDYBENSWSLSLCEV 540
Qy 880 DVCVPKATHQKEMDKISGKLESDSTSLSKILDTVHSCERARELQKDHCEQRTGKMEQMKKK 939
Db 541 DVCVPKATHQKEMDKISGKLESDSTSLSKILDTVHSCERARELQKDHCEQRTGKMEQMKKK 600
Qy 940 FCVLKXKLESEAKIQLSELENQKVKWEQLCSVR-LTL 975
Db 601 FCVLKXKLESEAKIQLSELENQKVKWEQLCSVRFLTL 637

RESULT 10
US-09-604-287A-469
; Sequence 469, Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (310)
; OTHER INFORMATION: Xaa = Any Amino Acid
; NAME/KEY: unsure
; LOCATION: (429)
; OTHER INFORMATION: Xaa = Any Amino Acid
; NAME/KEY: unsure
; LOCATION: (522)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-604-287A-469

Query Match 47.0%; Score 3252.5; DB 4; Length 650;
Best Local Similarity 98.9%; Pred. No. 7e-215;
Matches 630; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 340 MSPAKETSEKFTWAAGRPRIAKWEKETPVKTGCVARVTSNKTVKLEGRSKMIACPTK 399
Db 1 MSPAKETSEKFTWAAGRPRIAKWEKETPVKTGCVARVTSNKTVKLEGRSKMIACPTK 60
Qy 400 ESSTKASANDORPSESQDEDEYSCDSRLPSSAKIQVICIPESIQKMEINREVEE 459
Db 61 ESSTKASANDORPSESQDEDEYSCDSRLPSSAKIQVICIPESIQKMEINREVEE 120
Qy 460 PPKPSAFKPAIEMQNSVPNKAFELKNEOTLRADMPPEPSKQDYBENSWSLSLCEV 519
Db 121 PPKPSAFKPAIEMQNSVPNKAFELKNEOTLRADMPPEPSKQDYBENSWSLSLCEV 180
Qy 520 SQKDVCLPKATHQKEIDKINGKLESPNKGILLKATCGMKVSIPTKALELMDMTFFKAEP 579
Db 181 SQKDVCLPKATHQKEIDKINGKLESPNKGILLKATCGMKVSIPTKALELMDMTFFKAEP 240
Qy 580 PGKPSAFEPATEMOKSVNKALELKNQETWRADEILPSESQKQDYBENSWSLSLCEV 639
Db 241 PGKPSAFEPATEMOKSVNKALELKNQETWRADEILPSESQKQDYBENSWSLSLCEV 300
Qy 640 QKDVCLPKAAHQKEIDKINGKLESPVKGILLKANCCKVSIPTKALELMDMTFFKAEP 699
Db 301 QKDVCLPKAAHQKEIDKINGKLESPVKGILLKANCCKVSIPTKALELMDMTFFKAEP 360
```


Qy 700 EKPSAFPAIEMQKSVPNKALELKNQETLRADQMPFSESKQKDYESSWDSLSLCTVSQ 759
Db 361 EKPSAFPAIEMQKSVPNKALELKNQETLRADQMPFSESKQKDYESSWDSLSLCTVSQ 420
Qy 760 KDVCPLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMOTFKABPE 819
Db 421 KDVCPLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMOTFKABPE 480
Qy 820 KPSAFPAIEMQKSVPNKALELKNQETLRADQMPFSESKQKDYESSWDSLSLCTVSQ 879
Db 481 KPSAFPAIEMQKSVPNKALELKNQETLRADQMPFSESKQKDYESSWDSLSLCTVSQ 540
Qy 880 DVCVPKATHQKEMDKISGLESDSTLSKILDTVHSCERARELQKDHCEQRTGKMEQMKK 939
Db 541 DVCVPKATHQKEMDKISGLESDSTLSKILDTVHSCERARELQKDHCEQRTGKMEQMKK 600
Qy 940 FCVLKKLSEAKEIKSQLENOQKVKWEQBLCSVR-LTL 975
Db 601 FCVLKKLSEAKEIKSQLENOQKVKWEQBLCSVRFLTL 637

RESULT 11

US-09-834-759-469
; Sequence 469, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (310)
; OTHER INFORMATION: Xaa = Any Amino Acid
; NAME/KEY: unsure
; LOCATION: (429)
; OTHER INFORMATION: Xaa = Any Amino Acid
; NAME/KEY: unsure
; LOCATION: (522)
; OTHER INFORMATION: Xaa = Any Amino Acid

US-09-834-759-469

Query Match 47.0%; Score 3252.5; DB 4; Length 650;
Best Local Similarity 98.9%; Pred. No. 7e-215;
Matches 630; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
Qy 340 MSPAKETSEKFTWAAKGRPRKIAWEKKEPTVKTCGVARVTSNKTQVLEKGRSKMIACPTK 399
Db 1 MSPAKETSEKFTWAAKGRPRKIAWEKKEPTVKTCGVARVTSNKTQVLEKGRSKMIACPTK 60
Qy 400 ESSTKASANDQRPFSKQEEDEEYSCDSRSLFESSAKIQVCIPESIIYQKWEINREVEE 459
Db 61 ESSTKASANDQRPFSKQEEDEEYSCDSRSLFESSAKIQVCIPESIIYQKWEINREVEE 120
Qy 460 PPKPSAFKPAIEMQKSVPNKALELKNQETLRADQMPFSESKQKDYESSWDSLSLCTV 519
Db 121 PPKPSAFKPAIEMQKSVPNKALELKNQETLRADQMPFSESKQKDYESSWDSLSLCTV 180
Qy 520 SQKDVCPLPKATHQKEIDKINGKLEESPNKDGLLKATCGMKVSIPTKALELMDMOTFKABP 579

Db 181 SQKDVCPLPKATHQKEIDKINGKLEESPNKDGLLKATCGMKVSIPTKALELMDMOTFKABP 240
Qy 580 PGKPSAFEPATEMOKSVPNKALELKNQETLRADQMPFSESKQKDYESSWDSLSLCTVS 639
Db 241 PGKPSAFEPATEMOKSVPNKALELKNQETLRADQMPFSESKQKDYESSWDSLSLCTVS 300
Qy 640 QKDVCPLKAAHQKEIDKINGKLEESPVKDGILLKANCGMKVSIPTKALELMDMOTFKABPP 699
Db 301 QKDVCPLKAAHQKEIDKINGKLEESPVKDGILLKANCGMKVSIPTKALELMDMOTFKABPP 360
Qy 700 EKPSAFPAIEMQKSVPNKALELKNQETLRADQMPFSESKQKDYESSWDSLSLCTVSQ 759
Db 361 EKPSAFPAIEMQKSVPNKALELKNQETLRADQMPFSESKQKDYESSWDSLSLCTVSQ 420
Qy 760 KDVCPLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMOTFKABPE 819
Db 421 KDVCPLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMOTFKABPE 480
Qy 820 KPSAFPAIEMQKSVPNKALELKNQETLRADQMPFSESKQKDYESSWDSLSLCTVSQ 879
Db 481 KPSAFPAIEMQKSVPNKALELKNQETLRADQMPFSESKQKDYESSWDSLSLCTVSQ 540
Qy 880 DVCVPKATHQKEMDKISGLESDSTLSKILDTVHSCERARELQKDHCEQRTGKMEQMKK 939
Db 541 DVCVPKATHQKEMDKISGLESDSTLSKILDTVHSCERARELQKDHCEQRTGKMEQMKK 600
Qy 940 FCVLKKLSEAKEIKSQLENOQKVKWEQBLCSVR-LTL 975
Db 601 FCVLKKLSEAKEIKSQLENOQKVKWEQBLCSVRFLTL 637

RESULT 12

US-09-590-751A-469
; Sequence 469, Application US/09590751A
; Patent No. 6756477
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C6
; CURRENT APPLICATION NUMBER: US/09/590,751A
; CURRENT FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(650)
; OTHER INFORMATION: Xaa = Any Amino Acid

US-09-590-751A-469

Query Match 47.0%; Score 3252.5; DB 4; Length 650;
Best Local Similarity 98.9%; Pred. No. 7e-215;
Matches 630; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
Qy 340 MSPAKETSEKFTWAAKGRPRKIAWEKKEPTVKTCGVARVTSNKTQVLEKGRSKMIACPTK 399
Db 1 MSPAKETSEKFTWAAKGRPRKIAWEKKEPTVKTCGVARVTSNKTQVLEKGRSKMIACPTK 60
Qy 400 ESSTKASANDQRPFSKQEEDEEYSCDSRSLFESSAKIQVCIPESIIYQKWEINREVEE 459
Db 61 ESSTKASANDQRPFSKQEEDEEYSCDSRSLFESSAKIQVCIPESIIYQKWEINREVEE 120
Qy 460 PPKPSAFKPAIEMQKSVPNKALELKNQETLRADQMPFSESKQKDYESSWDSLSLCTV 519

Db 121 PPKPSAFKPAIEMQNSVPNKAPELKNQOTLRADPMFPPESSKQKDYENSWDSLSLCTV 180
Qy 520 SQKDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDQMTFKAEP 579
Db 181 SQKDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDQMTFKAEP 240
Qy 580 PKPSAFPEPATMOKSVNPKALELKNQOTLRADPMFPPESSKQKDYENSWDSLSLCTV 639
Db 241 PKPSAFPEPATMOKSVNPKALELKNQOTLRADPMFPPESSKQKDYENSWDSLSLCTV 300
Qy 640 QKDVCLPKAAHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDQMTFKAEP 699
Db 301 QKDVCLPKAAHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDQMTFKAEP 360
Qy 700 EKPSAFPEPATMOKSVNPKALELKNQOTLRADPMFPPESSKQKDYENSWDSLSLCTV 759
Db 361 EKPSAFPEPATMOKSVNPKALELKNQOTLRADPMFPPESSKQKDYENSWDSLSLCTV 420
Qy 760 KDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDQMTFKAEP 819
Db 421 KDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDQMTFKAEP 480
Qy 820 KPSAFPEPATMOKSVNPKALELKNQOTLRADPMFPPESSKQKDYENSWDSLSLCTV 879
Db 481 KPSAFPEPATMOKSVNPKALELKNQOTLRADPMFPPESSKQKDYENSWDSLSLCTV 540
Qy 880 DVCVPKATHQKEMDKISGKLEDSTLSKILDTVHSCERARELQKDHCEQRTGKMEQMKK 939
Db 541 DVCVPKATHQKEMDKISGKLEDSTLSKILDTVHSCERARELQKDHCEQRTGKMEQMKK 600
Qy 940 FCVLKXKLSEAKEIKSQLENQKVKWEQELCSVR-LTL 975
Db 601 FCVLKXKLSEAKEIKSQLENQKVKWEQELCSVRFLTL 637

RESULT 13

US-09-551-621-469.
; Sequence 469, Application US/09551621
; Patent No. 6825175
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C5
; CURRENT APPLICATION NUMBER: US/09/551,621
; CURRENT FILING DATE: 2000-04-17
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (310)
; NAME/KEY: Xaa = Any Amino Acid<221> unsure
; LOCATION: (429)
; NAME/KEY: Xaa = Any Amino Acid<221> unsure
; LOCATION: (522)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-551-621-469

Query Match 47.0%; Score 3252.5; DB 4; Length 650;
Best Local Similarity 98.9%; Pred. No. 7e-215;
Matches 630; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 340 MSPAKETSEKFTAAAGPRKIAPKKEPTVKTCGVARTSNKTKVLEKGRSKMIACPTK 399
Db 1 MSPAKETSEKFTAAAGPRKIAPKKEPTVKTCGVARTSNKTKVLEKGRSKMIACPTK 60

Qy 400 ESSTKASANDORFPSEKQEEDEEYSCDSRSLFESSAKIQVCIPESIYQKVMENREVEE 459
Db 61 ESSTKASANDORFPSEKQEEDEEYSCDSRSLFESSAKIQVCIPESIYQKVMENREVEE 120
Qy 460 PPKPSAFKPAIEMQNSVPNKAPELKNQOTLRADPMFPPESSKQKDYENSWDSLSLCTV 519
Db 121 PPKPSAFKPAIEMQNSVPNKAPELKNQOTLRADPMFPPESSKQKDYENSWDSLSLCTV 180
Qy 520 SQKDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDQMTFKAEP 579
Db 181 SQKDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDQMTFKAEP 240
Qy 580 PKPSAFPEPATMOKSVNPKALELKNQOTLRADPMFPPESSKQKDYENSWDSLSLCTV 639
Db 241 PKPSAFPEPATMOKSVNPKALELKNQOTLRADPMFPPESSKQKDYENSWDSLSLCTV 300
Qy 640 QKDVCLPKAAHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDQMTFKAEP 699
Db 301 QKDVCLPKAAHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDQMTFKAEP 360
Qy 700 EKPSAFPEPATMOKSVNPKALELKNQOTLRADPMFPPESSKQKDYENSWDSLSLCTV 759
Db 361 EKPSAFPEPATMOKSVNPKALELKNQOTLRADPMFPPESSKQKDYENSWDSLSLCTV 420
Qy 760 KDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDQMTFKAEP 819
Db 421 KDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDQMTFKAEP 480
Qy 820 KPSAFPEPATMOKSVNPKALELKNQOTLRADPMFPPESSKQKDYENSWDSLSLCTV 879
Db 481 KPSAFPEPATMOKSVNPKALELKNQOTLRADPMFPPESSKQKDYENSWDSLSLCTV 540
Qy 880 DVCVPKATHQKEMDKISGKLEDSTLSKILDTVHSCERARELQKDHCEQRTGKMEQMKK 939
Db 541 DVCVPKATHQKEMDKISGKLEDSTLSKILDTVHSCERARELQKDHCEQRTGKMEQMKK 600
Qy 940 FCVLKXKLSEAKEIKSQLENQKVKWEQELCSVR-LTL 975
Db 601 FCVLKXKLSEAKEIKSQLENQKVKWEQELCSVRFLTL 637

RESULT 14

US-09-620-405B-494
; Sequence 494, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 494
; LENGTH: 743
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(743)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-620-405B-494

Query Match 47.0%; Score 3252.5; DB 4; Length 743;
Best Local Similarity 98.9%; Pred. No. 8.5e-215;
Matches 630; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:52:33 ; Search time 146.087 Seconds
(without alignments)
3550.262 Million cell updates/sec

Title: US-09-602-362E-23

Perfect score: 6320

Sequence: 1 MTRKKKTINLIQAQRKA.....NHLKNRIYQVEKEKAETENS 1341

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6920	100.0	1341	4	AAB84702
2	6920	100.0	1341	5	Abj05537 Breast ca
3	6920	100.0	1341	6	Abj37784 Human tum
4	6920	100.0	1341	6	Abj47548 Breast ca
5	6920	100.0	1341	7	Adl93227 Human bre
6	6892	99.6	1349	6	Abj37788 Human tum
7	6892	99.6	1349	7	Adl93235 Human bre
8	5130	74.1	1002	4	AAU33351 Human bre
9	5130	74.1	1002	5	Abg78918 Human bre
10	5130	74.1	1002	6	Abj37741 Human tum
11	5130	74.1	1002	7	Adl93137 Human bre
12	5130	74.1	1002	8	Adl93137 Human bre
13	5130	74.1	1095	4	AAU33357 Human bre
14	5130	74.1	1095	5	Abg78924 Human bre
15	5130	74.1	1095	6	Abj37747 Human tum
16	5130	74.1	1095	7	Adl93155 Human bre
17	5124	74.0	1013	6	Abj37783 Human tum
18	5124	74.0	1013	7	Adl93215 Human bre
19	3945	57.0	1239	6	Abj37789 Human tum
20	3945	57.0	1239	7	Adl93239 Human bre
21	3721	53.8	1225	6	Abj47547 Breast ca
22	3256.5	47.1	661	6	Abj37782 Human tum
23	3256.5	47.1	661	7	Adl93214 Human bre
24	3252.5	47.0	650	4	AAAB50263 Human bre
25	3252.5	47.0	650	4	AAAG65983 B726P spl

ALIGNMENTS

RESULT 1

AAAB84702
ID AAB84702 standard; protein; 1341 AA.

XX AAB84702;

AC AAB84702;

DT 17-SEP-2001 (first entry)

DE Amino acid sequence of a human cancer associated antigen.

XX Cancer associated antigen; INGI; tumour suppressor; cancer; vaccine.

OS Homo sapiens.

XX WO200147959-A2.

PD 05-JUL-2001.

XX 29-NOV-2000; 2000WO-US042334.

XX 30-NOV-1999; 99US-00451739.

PR 24-OCT-2000; 2000US-00602362.

XX (LUDW-) LUDWIG INST CANCER RES.

XX (SLOK) SLOAN KETTERING INST CANCER RES.

XX (CORR) CORNELL RES FOUND INC.

PI Jager D, Stockert E, Scanlan M, Knuth A, Old L, Gure A, Chen Y;

XX WPI; 2001-441706/47.

XX Isolated cancer associated nucleic acid molecule identified by SEREX

XX (serological identification of antigens by recombinant expression

XX cloning) technique, useful in nucleic acid based therapies to treat

XX cancer.

XX Claim 83; Page 53-57; 62pp; English.

XX The present sequence represents a human cancer associated antigen. The

XX sequence was identified using probes derived from the INGI gene. The INGI

XX gene is a tumour suppressor candidate gene. The cancer associated antigen

XX polynucleotides and polypeptides are useful for screening for the

XX possible presence of a pathological condition in a subject such as

AAU33346 Human bre
Abg78913 Human bre
Abj37736 Human tum
Adl93131 Human bre
Ade44421 Human bre
AAU33358 Human bre
Abg78925 Human bre
Abj37748 Human tum
Adl93156 Human bre
Aab84703 Amino aci
Adj69751 Human hea
Aab84701 Amino aci
AAB50248 Human bre
AAG65986 B726P spl
AAU33349 Human bre
Abj37739 Human tum
Adl93134 Human bre
Ade44424 Human bre
AAB50249 Human bre

Query Match 100.0%; Score 6920; DB 4; Length 1341;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTKRKTINLNIQDAQKTAHLWACVNGHEEVVTFVLDVKQDQVLDGEHRTPLMKALQC 60
DB 1 MTKRKTINLNIQDAQKTAHLWACVNGHEEVVTFVLDVKQDQVLDGEHRTPLMKALQC 60

QY 61 HOECANLIDSGADINLDVYGNMALHYAVYSEILSVAKLISHGAVIEVHNKASLTPL 120
DB 61 HOECANLIDSGADINLDVYGNMALHYAVYSEILSVAKLISHGAVIEVHNKASLTPL 120

QY 121 LLSITRSEQIVFLLIKNANANAVNKYKTALMLAVCHGSSEIVGMLLQNVDFVFAADI 180
DB 121 LLSITRSEQIVFLLIKNANANAVNKYKTALMLAVCHGSSEIVGMLLQNVDFVFAADI 180

QY 181 CGVTAHYAVTCGFFHHIEQIMYIRKLSKNHNTNPEGSTAGTPDEAAPLAERTPTDAE 240
DB 181 CGVTAHYAVTCGFFHHIEQIMYIRKLSKNHNTNPEGSTAGTPDEAAPLAERTPTDAE 240

QY 241 SLVEKTPDEAAPLVBERTPDTAESLVEKTPDEAASLVEGTSKIOCLEKATSGKFEQSAEE 300
DB 241 SLVEKTPDEAAPLVBERTPDTAESLVEKTPDEAASLVEGTSKIOCLEKATSGKFEQSAEE 300

QY 301 TPREITSPAKETSEKFTWPAKGRPKRIAEWKEDTPREIMSPAKETSEKFTWAAKGRPRK 360
DB 301 TPREITSPAKETSEKFTWPAKGRPKRIAEWKEDTPREIMSPAKETSEKFTWAAKGRPRK 360

QY 361 IAWKKEPTVTKGCVARVTSNKTIVLEKGRSKMIACPTKESSTKASANDORPPSSKQEE 420
DB 361 IAWKKEPTVTKGCVARVTSNKTIVLEKGRSKMIACPTKESSTKASANDORPPSSKQEE 420

QY 421 DEEYSCDSRSLFESSAKIOVCIPEISYOKVMEINREVEEPKPSAFPAIEMQNSVFNK 480
DB 421 DEEYSCDSRSLFESSAKIOVCIPEISYOKVMEINREVEEPKPSAFPAIEMQNSVFNK 480

QY 481 APELKNEQTLRADPMFPPESSKQDYENSWSDESICETVSQKDVCLPKATHQKEIDKING 540
DB 481 APELKNEQTLRADPMFPPESSKQDYENSWSDESICETVSQKDVCLPKATHQKEIDKING 540

QY 541 KLEESPNKDGILLKATCGMKVSIPTKALEIKDMQTFKAEPPKPSAFEPATEMOKSVPNKA 600
DB 541 KLEESPNKDGILLKATCGMKVSIPTKALEIKDMQTFKAEPPKPSAFEPATEMOKSVPNKA 600

QY 601 LELKNEQTLRADEILPSSKQDYENSWSDESICETVSQKDVCLPKAAHOKEDIDKING 660
DB 601 LELKNEQTLRADEILPSSKQDYENSWSDESICETVSQKDVCLPKAAHOKEDIDKING 660

QY 661 LEGSPVKDGLLKANCCKMVSIPTKALELMDMTQTFKAEPPKPSAFEPATEMOKSVPNKAL 720
DB 661 LEGSPVKDGLLKANCCKMVSIPTKALELMDMTQTFKAEPPKPSAFEPATEMOKSVPNKAL 720

QY 721 ELKNEQTLRADEILPSSKQDYENSWSDESICETVSQKDVCLPKATHQKEIDKINGKL 780
DB 721 ELKNEQTLRADEILPSSKQDYENSWSDESICETVSQKDVCLPKATHQKEIDKINGKL 780

QY 781 BESPNDCGLKAPCRMKVSIPTKALELMDMTQTFKAEPPKPSAFEPATEMOKSVPNKALE 840
DB 781 BESPNDCGLKAPCRMKVSIPTKALELMDMTQTFKAEPPKPSAFEPATEMOKSVPNKALE 840

QY 841 LKNEQTLRADQMFPPSSKQDYENSWSDESILRETVSQKDVCPKATHQKEMDKISGKLE 900
DB 841 LKNEQTLRADQMFPPSSKQDYENSWSDESILRETVSQKDVCPKATHQKEMDKISGKLE 900

QY 901 DSTLSLSKILDVHSCERARELOKHCEORTGMEQMKKFCVCLKKSEAKEIKSOLNQ 960
DB 901 DSTLSLSKILDVHSCERARELOKHCEORTGMEQMKKFCVCLKKSEAKEIKSOLNQ 960

QY 961 KVKWEQELCSVRLTLNQEEKERNADIINEKIREELGRIEEHOHRELEVKQOLEQALRIQ 1020
DB 961 KVKWEQELCSVRLTLNQEEKERNADIINEKIREELGRIEEHOHRELEVKQOLEQALRIQ 1020

QY 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDI 1080

DB 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDI 1080
QY 1081 KILKEKNAELQMTLKLKEESLTKEASQYSGOLKVLIAENTMTLSKLKEKODKEILEASIE 1140
DB 1081 KILKEKNAELQMTLKLKEESLTKEASQYSGOLKVLIAENTMTLSKLKEKODKEILEASIE 1140

QY 1141 SHHPLASAVODHQIIVTSRKSQEPAPHIAAGDACLQKMNVDVSVSTIYNNVHLHOPLSEA 1200
DB 1141 SHHPLASAVODHQIIVTSRKSQEPAPHIAAGDACLQKMNVDVSVSTIYNNVHLHOPLSEA 1200

QY 1201 QRSKSLKININLYAGDALRENTLVSEHAQRDQRETQCOMKEAHHYQNEQDNVKNHTSQ 1260
DB 1201 QRSKSLKININLYAGDALRENTLVSEHAQRDQRETQCOMKEAHHYQNEQDNVKNHTSQ 1260

QY 1261 ESLDQKLFQLOSKNMWLQOQLVHAHKADNKSKITIDHFLERKMOHLLKKEKEIEIFNY 1320
DB 1261 ESLDQKLFQLOSKNMWLQOQLVHAHKADNKSKITIDHFLERKMOHLLKKEKEIEIFNY 1320

QY 1321 NNHLKNRIYQYEKEKAETENS 1341
DB 1321 NNHLKNRIYQYEKEKAETENS 1341

RESULT 2
ABJ05537
ID ABJ05537 standard; protein; 1341 AA.
XX
AC ABJ05537;
XX
DT 14-NOV-2002 (first entry)
XX
DE Breast cancer-associated protein 2.
XX
KW Breast cancer; breast cancer-associated gene sequence; drug development;
XX pharmacogenetics; biosensor development.
XX Unidentified.
XX WO200259377-A2.
XX
PD 01-AUG-2002.
XX
PF 24-JAN-2002; 2002WO-US002242;
XX
PR 24-JAN-2001; 2001US-0263965P.
PR 02-FEB-2001; 2001US-0265928P.
PR 09-APR-2001; 2001US-00829472.
PR 09-APR-2001; 2001US-0282698P.
PR 04-MAY-2001; 2001US-0288590P.
PR 29-MAY-2001; 2001US-0294443P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Mack DH, Gish KC, Afar D;
XX
WPI; 2002-583738/62.
XX N-PSDB; ABT07694.
PT Detecting a breast cancer-associated transcript in a patient's cell,
PT useful for diagnosing breast cancer, comprises contacting a biological
PT sample with a polynucleotide that selectively hybridizes with breast
PT cancer nucleic acids.
XX
PS Disclosure; Page 348-349; 414pp; English.
XX
CC The invention comprises a method of detecting a breast cancer-associated
CC transcript in a cell from a patient. The method of the invention involves
CC contacting a biological sample from the patient with a nucleotide that
CC hybridises to one of the 69 breast cancer-associated gene sequences shown
CC in the specification. The method of the invention is useful in the
CC diagnosis or prognosis of breast cancer, and for detecting genes that are
CC up or down-regulated in breast cancer cells. Genes identified by the

method of the invention can be used in diagnostic purposes and also as targets for screening for therapeutic compounds that modulate breast cancer (e.g. hormones or antibodies). Identification of genes that are over or under expressed in breast cancer can additionally provide high-resolution, high-sensitivity datasets which can be used in the areas of diagnostics, therapeutics, drug development, pharmacogenetics, protein structure and biosensor development. Amino acid sequences ABJ05536 - ABJ05604 represent the proteins encoded by the 69 breast cancer-associated genes of the invention

CC SQ Sequence 1341 AA;

Query Match 100.0%; Score 6920; DB 5; Length 1341;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTKKKKTININIQDAQKRTALHWACVNGCHERVVTLVDKCOLDVLGDEHRTPLMKALQC 60
Db 1 MTKKKKTININIQDAQKRTALHWACVNGCHEEVVTLVDKCOLDVLGDEHRTPLMKALQC 60
Qy 61 HQEACANILDSGADINLVDVYGNMALHYAVYSILSVAKLLSHGAVIEVHNKASLTPL 120
Db 61 HQEACANILDSGADINLVDVYGNMALHYAVYSILSVAKLLSHGAVIEVHNKASLTPL 120
Qy 121 LLSITKRSQIPEVFLLIKNNANANVKKCTALMLAVCHGSSEIVGMLLQONVDFPAADI 180
Db 121 LLSITKRSQIPEVFLLIKNNANANVKKCTALMLAVCHGSSEIVGMLLQONVDFPAADI 180
Qy 181 CGVTAHYAVTCGPHHHEQIMEYIRKLSKNHONTNPEGTSAGTPDEAAPLAERTPDTA 240
Db 181 CGVTAHYAVTCGPHHHEQIMEYIRKLSKNHONTNPEGTSAGTPDEAAPLAERTPDTA 240
Qy 241 SLVEKTPDEAAPLVERTPDAESLVEKTPDAESLVEKTPDAESLVEKTPDAESLVEKTPDA 300
Db 241 SLVEKTPDEAAPLVERTPDAESLVEKTPDAESLVEKTPDAESLVEKTPDAESLVEKTPDA 300
Qy 301 TPREITSPAKETSEKFTWPAKGRPRKIAWEKEDTPREIMSPAKETSEKFTWPAKGRPRK 360
Db 301 TPREITSPAKETSEKFTWPAKGRPRKIAWEKEDTPREIMSPAKETSEKFTWPAKGRPRK 360
Qy 361 IAWKKETPVKTCVAVTSNKTIVLEKGRSKMTACPTKESSTKASANDORFPSESQOE 420
Db 361 IAWKKETPVKTCVAVTSNKTIVLEKGRSKMTACPTKESSTKASANDORFPSESQOE 420
Qy 421 DEEYSCDSRSLFESSAKIQVCPESIQVKWMEINREVEEPPKPSAFKPAIEMQNSVPNK 480
Db 421 DEEYSCDSRSLFESSAKIQVCPESIQVKWMEINREVEEPPKPSAFKPAIEMQNSVPNK 480
Qy 481 APELKNQTLRADPMFPPEKQKDYBENSWSLCEVTSQKDVCLPKATHQKEIDKING 540
Db 481 APELKNQTLRADPMFPPEKQKDYBENSWSLCEVTSQKDVCLPKATHQKEIDKING 540
Qy 541 KLEESPNKDGLLKATCGMKVSIPTKALELMDMTFKAEPPGKPSAFEPATEMOKSVNKA 600
Db 541 KLEESPNKDGLLKATCGMKVSIPTKALELMDMTFKAEPPGKPSAFEPATEMOKSVNKA 600
Qy 601 LELKNEOTWRADEILPESKQKDYBENSWSLCEVTSQKDVCLPKAAHOKETDKINGK 660
Db 601 LELKNEOTWRADEILPESKQKDYBENSWSLCEVTSQKDVCLPKAAHOKETDKINGK 660
Qy 661 LEGSPVKDGLLKANCGRMKVSIPTKALELMDMTFKAEPPPEKPSAFEPATEMOKSVPNKAL 720
Db 661 LEGSPVKDGLLKANCGRMKVSIPTKALELMDMTFKAEPPPEKPSAFEPATEMOKSVPNKAL 720
Qy 721 ELKNEOTLRADEILPESKQKDYBENSWSLCEVTSQKDVCLPKATHQKEIDKINGKL 780
Db 721 ELKNEOTLRADEILPESKQKDYBENSWSLCEVTSQKDVCLPKATHQKEIDKINGKL 780
Qy 781 EESPDNDGFLKAPCRMVSIPTKALELMDMTFKAEPPKPSAFEPATEMOKSVPNKALE 840
Db 781 EESPDNDGFLKAPCRMVSIPTKALELMDMTFKAEPPKPSAFEPATEMOKSVPNKALE 840
Qy 841 LKNEQTLRADQMFPPESEKQKDYBENSWSLCEVTSQKDVCPKATHQKEMDKISGKLE 900

Db 841 LKNEQTLRADQMFPPESEKQKDYBENSWSLCEVTSQKDVCPKATHQKEMDKISGKLE 900
Qy 901 DSTSLSKILDTVHSCERARELQKDHCHORTGKMEQMKKFCVLKKKLSEAKEIKSOLBNQ 960
Db 901 DSTSLSKILDTVHSCERARELQKDHCHORTGKMEQMKKFCVLKKKLSEAKEIKSOLBNQ 960
Qy 961 KVKWEQELCSVRLTLNQEEBKKRRNADILNEKIREBELGRIBEQHRKELEVKQOLEQALRIQ 1020
Db 961 KVKWEQELCSVRLTLNQEEBKKRRNADILNEKIREBELGRIBEQHRKELEVKQOLEQALRIQ 1020
Qy 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKKIAMLKLEIATLKHOYQEKENKYFEDI 1080
Db 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKKIAMLKLEIATLKHOYQEKENKYFEDI 1080
Qy 1081 KILKEKNAELQMTLKLKEESLTKRASQVSGOLKVLIAENTMLTSLKLEKQDKLELEABIE 1140
Db 1081 KILKEKNAELQMTLKLKEESLTKRASQVSGOLKVLIAENTMLTSLKLEKQDKLELEABIE 1140
Qy 1141 SHHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQRKMNVDVSTIYNNEVLHQPSEA 1200
Db 1141 SHHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQRKMNVDVSTIYNNEVLHQPSEA 1200
Qy 1201 QRKSKSLKINLYAGDALRENTLVSEHAORDORETQCOMKEAHEMYQNEODNVNKHTEQQ 1260
Db 1201 QRKSKSLKINLYAGDALRENTLVSEHAORDORETQCOMKEAHEMYQNEODNVNKHTEQQ 1260
Qy 1261 ESLDQKLPQLQSKNMVLQOQLVHAHKADNKSKITIDHFLERKMQHLLKKEKNEEIEFNY 1320
Db 1261 ESLDQKLPQLQSKNMVLQOQLVHAHKADNKSKITIDHFLERKMQHLLKKEKNEEIEFNY 1320
Qy 1321 NNHLKNRIYQYEKEKAETENS 1341
Db 1321 NNHLKNRIYQYEKEKAETENS 1341
RESULT 3
ABJ37784
ID ABJ37784 standard; protein; 1341 AA.
XX AC ABJ37784;
XX DT 15-MAY-2003 (first entry)
XX DE Human tumour-related protein - SEQ ID No 565.
XX KW Human; vaccine; gene therapy; T cell stimulation; T cell expansion;
XX KW tumour; breast cancer; cancer; immune response stimulation.
XX OS Homo sapiens.
XX PN WO200283956-A1.
XX PD 24-OCT-2002.
XX PF 15-APR-2002; 2002WO-US012378.
XX PR 13-APR-2001; 2001US-00834759.
XX PR 07-DEC-2001; 2001US-00007805.
XX PR 13-FEB-2002; 2002US-00076622.
XX PA (CORI-) CORIXA CORP.
XX PI Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;
PI Vedvick TS, McNeill PD, Durham M;
XX WPI; 2003-103376/09.
XX PT New polypeptide and polynucleotide useful for stimulating and/or
XX PT expanding T cells specific for a tumor protein and treating breast
XX PT cancer.

PS Example 9; Page 342-346; 375pp; English.

XX The invention comprises a method of stimulating and/or expanding T cells specific for a tumour protein. The invention further comprises human CC nucleic acids and proteins that are associated with tumours (e.g. breast CC cancer). The method and sequences of the invention are useful for CC stimulating and/or expanding T cells specific for a tumour protein, CC detecting the presence of cancer, stimulating an immune response in a CC patient and treating breast cancer. The present amino acid sequence CC represents a human tumour-related protein

XX Sequence 1341 AA;

Query Match 100.0%; Score 6920; DB 6; Length 1341;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTKRKTINLNIQDAKRTALHWACVNGHEEVTFLVDRKQCLDVLGDEHRTPLMKALQC 60

Db 1 MTKRKTINLNIQDAKRTALHWACVNGHEEVTFLVDRKQCLDVLGDEHRTPLMKALQC 60

QY 61 HOEACANILDSGADINLVGYGNMAYVYSEILSVVAKLHSGAVIEVHNKASLTPL 120

Db 61 HOEACANILDSGADINLVGYGNMAYVYSEILSVVAKLHSGAVIEVHNKASLTPL 120

QY 121 LLSITKRSQIPEFLLIKNANAVNKYCTALMLAVCHGSSEIVGMLLQONVDVFAADI 180

Db 121 LLSITKRSQIPEFLLIKNANAVNKYCTALMLAVCHGSSEIVGMLLQONVDVFAADI 180

QY 181 CGVTAHVAVTCGPHHIEQIMYIRKLSKNHNTNPEGSTAGTPDEAALPRTPTAE 240

Db 181 CGVTAHVAVTCGPHHIEQIMYIRKLSKNHNTNPEGSTAGTPDEAALPRTPTAE 240

QY 241 SLVEKTPDEAALPRTPTAEISLVEKTPDEAASLVEGTSKIOCLEKATSGKFEQSABE 300

Db 241 SLVEKTPDEAALPRTPTAEISLVEKTPDEAASLVEGTSKIOCLEKATSGKFEQSABE 300

QY 301 TPREITSPAKTSEKFTWPAKGRPKIAWEKEDTPREIMSPAKTSEKFTWAAKGRPK 360

Db 301 TPREITSPAKTSEKFTWPAKGRPKIAWEKEDTPREIMSPAKTSEKFTWAAKGRPK 360

QY 361 IAWKKEPTVKTCVARTSNKTKVLEKGRSKMIACPTKESSTKASANDORPPSESKEE 420

Db 361 IAWKKEPTVKTCVARTSNKTKVLEKGRSKMIACPTKESSTKASANDORPPSESKEE 420

QY 421 DEEYSCDSRSLPSSAKIQVCIPESYIYQKWEINREVEEPPKPSAPKPAIEMQNSVFNK 480

Db 421 DEEYSCDSRSLPSSAKIQVCIPESYIYQKWEINREVEEPPKPSAPKPAIEMQNSVFNK 480

QY 481 APELNEQTLRADPMPFPESKQDYENSWSDESICETVSKQDVCLPKATHQKEIDKING 540

Db 481 APELNEQTLRADPMPFPESKQDYENSWSDESICETVSKQDVCLPKATHQKEIDKING 540

QY 541 KLEESPNKDGLLKATCGMKVSIPTKALELMDMTFKAEPPGKPSAFEPATEMOKSVNKA 600

Db 541 KLEESPNKDGLLKATCGMKVSIPTKALELMDMTFKAEPPGKPSAFEPATEMOKSVNKA 600

QY 601 LELKNEQTLRADPMPFPESKQDYENSWSDESICETVSKQDVCLPKAAHQKEIDKING 660

Db 601 LELKNEQTLRADPMPFPESKQDYENSWSDESICETVSKQDVCLPKAAHQKEIDKING 660

QY 661 LEGSPVKDGLKANCCKVSIPTKALELMDMTFKAEPPKPSAFEPATEMOKSVNKA 720

Db 661 LEGSPVKDGLKANCCKVSIPTKALELMDMTFKAEPPKPSAFEPATEMOKSVNKA 720

QY 721 ELKNEQTLRADPMPFPESKQDYENSWSDESICETVSKQDVCLPKATHQKEIDKING 780

Db 721 ELKNEQTLRADPMPFPESKQDYENSWSDESICETVSKQDVCLPKATHQKEIDKING 780

QY 781 EESPNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPKPSAFEPATEMOKSVNKA 840

Db 781 EESPNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPKPSAFEPATEMOKSVNKA 840

QY 841 LKNEQTLRADOMPFPESKQDYENSWSDESICETVSKQDVCPKATHQKEMDKISGLE 900

Db 841 LKNEQTLRADOMPFPESKQDYENSWSDESICETVSKQDVCPKATHQKEMDKISGLE 900

QY 901 DSTSLKILDTVHSCBRARELOKHCEORTKQWQMKKFCVLKKLSEAKEIKSQLENO 960

Db 901 DSTSLKILDTVHSCBRARELOKHCEORTKQWQMKKFCVLKKLSEAKEIKSQLENO 960

QY 961 KVKWEQELCSVRILTNOEERKRNADILNEKIRIEELGRIEEQHRLKEVQKQLEQALRIQ 1020

Db 961 KVKWEQELCSVRILTNOEERKRNADILNEKIRIEELGRIEEQHRLKEVQKQLEQALRIQ 1020

QY 1021 DIELKSVEINLQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHOYQEKENKYPEDI 1080

Db 1021 DIELKSVEINLQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHOYQEKENKYPEDI 1080

QY 1081 KILKEKNAELQMTLKLKESLITKASQYSQQLVLAENTMLTSLKLEKQDKETLEAEIE 1140

Db 1081 KILKEKNAELQMTLKLKESLITKASQYSQQLVLAENTMLTSLKLEKQDKETLEAEIE 1140

QY 1141 SHHPRLASAVODHDOIVTSRKSOEPAPHAGDACLQRMNVDSSTIYNNVHLQPLSEA 1200

Db 1141 SHHPRLASAVODHDOIVTSRKSOEPAPHAGDACLQRMNVDSSTIYNNVHLQPLSEA 1200

QY 1201 QRKSKSLKINLVAGDALRENTLVSEHAQDQRETOCQMKAEAHMYQNEQDNVNHKTEQQ 1260

Db 1201 QRKSKSLKINLVAGDALRENTLVSEHAQDQRETOCQMKAEAHMYQNEQDNVNHKTEQQ 1260

QY 1261 ESLDQKLFLQSKNMWLOQLVHAHKADNKSITTDIHFLEKMKOHHLKKNKEEIFNY 1320

Db 1261 ESLDQKLFLQSKNMWLOQLVHAHKADNKSITTDIHFLEKMKOHHLKKNKEEIFNY 1320

QY 1321 NNHLKNRIYQYEKEASTENS 1341

Db 1321 NNHLKNRIYQYEKEASTENS 1341

RESULT 4

ABR47548

ID ABR47548 standard; protein; 1341 AA.

XX AC ABR47548;

XX XX

DT 12-JUN-2003 (first entry)

DE Breast cancer associated protein sequence SEQ ID NO:334.

XX KW Human; breast cancer; cytostatic; gene therapy.

XX OS Homo sapiens.

XX XX

PN WO2003004989-A2.

XX PD 16-JAN-2003.

XX PF 21-JUN-2002; 2002WO-US019669.

XX PR 21-JUN-2001; 2001US-0299887P.

PR 27-JUN-2001; 2001US-0301572P.

PR 18-JUL-2001; 2001US-0306501P.

PR 25-SEP-2001; 2001US-0325002P.

PR 05-MAR-2002; 2002US-0362585P.

PR 14-MAY-2002; 2002US-0380391P.

XX (MILL-) MILLENIUM PHARM INC.

XX Lillie J, Gannavarapu M, Glatt K, Hoersh S, Kamatkar S;

PI Mertens M, Monahan JE, Myer Y, Wang Y, Xu Y, Zhao X, Meyers RE;

XX Bast RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;

DR WPI: 2003-210381/20.

XX N-PSDB; ACC50246.

PT Breast cancer diagnosis or treatment by comparing the level of expression
 PT of a marker in a patient sample with that in the control non-breast
 PT cancer sample.

XX Claim 1; SEQ ID NO 334; 128pp; English.

XX The present invention describes a method for assessing whether a patient
 CC is afflicted with breast cancer. The method comprises comparing the level
 CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
 CC ABR47386 to ABR47632) in a patient sample and the normal level of
 CC expression of the marker in a control non-breast cancer sample, where a
 CC significant increase in the level of expression of the marker in the
 CC patient sample and the normal level is an indication that the patient is
 CC afflicted with breast cancer. The breast cancer associated sequences from
 CC the present invention have cytostatic activities and can be used in gene
 CC therapy. The method is useful for diagnosing and treating breast cancer.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1341 AA;

Query Match 100.0%; Score 6920; DB 6; Length 1341;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTKKKTINLNIOQAQRKTHLWACVNGHEEVVTLVDRKQDLVDLGEHRTPLMKALQC 60
 Db 1 MTKKKTINLNIOQAQRKTHLWACVNGHEEVVTLVDRKQDLVDLGEHRTPLMKALQC 60

Qy 61 HQEACANILDSGADINLVGYGNALHYAVYSILSVAKLLSHGAVIEVHNKASLTPL 120
 Db 61 HQEACANILDSGADINLVGYGNALHYAVYSILSVAKLLSHGAVIEVHNKASLTPL 120

Qy 121 LLSITKSEQIVTELLIKNANANAVKYKCTALMLAVCHSGSEIVGMLLQNVDFPAADI 180
 Db 121 LLSITKSEQIVTELLIKNANANAVKYKCTALMLAVCHSGSEIVGMLLQNVDFPAADI 180

Qy 181 CGVTAHYAVTCGPHHIEQIMEYIRKLSKNHONTNPEGTSAGTPDEAALPAERTPTDAE 240
 Db 181 CGVTAHYAVTCGPHHIEQIMEYIRKLSKNHONTNPEGTSAGTPDEAALPAERTPTDAE 240

Qy 241 SLVEKTPDEAALVETPTDAESLVEKTPDEAASLVGTSVKIQCLEKATSGKFEQSAEE 300
 Db 241 SLVEKTPDEAALVETPTDAESLVEKTPDEAASLVGTSVKIQCLEKATSGKFEQSAEE 300

Qy 301 TPREITSPAKETSEKFTWPAKGRPRKTAWEKEDTPREIMSPAKETSEKFTWAAKGRPRK 360
 Db 301 TPREITSPAKETSEKFTWPAKGRPRKTAWEKEDTPREIMSPAKETSEKFTWAAKGRPRK 360

Qy 361 IAWKKEKTPVKTGCVARVTSNKTVLEKGRSKMIACPTKESSTKASANDORPPSESQKEE 420
 Db 361 IAWKKEKTPVKTGCVARVTSNKTVLEKGRSKMIACPTKESSTKASANDORPPSESQKEE 420

Qy 421 DEEYSCDSRSLFESSAKIQVCPISIIYQKWEINREVEEPPKPSAFKPAIEMQNSVPNK 480
 Db 421 DEEYSCDSRSLFESSAKIQVCPISIIYQKWEINREVEEPPKPSAFKPAIEMQNSVPNK 480

Qy 481 APELKNQOTLRADMPFPESKQDYBESSWDSLSLCEVTSQKDVCLPKATHQKEIDKING 540
 Db 481 APELKNQOTLRADMPFPESKQDYBESSWDSLSLCEVTSQKDVCLPKATHQKEIDKING 540

Qy 541 KLEBSPNKGILLKATCGMKVSIPTKALELMDQMTFKAEPGKPSAFEPATEMOKSVPNKA 600
 Db 541 KLEBSPNKGILLKATCGMKVSIPTKALELMDQMTFKAEPGKPSAFEPATEMOKSVPNKA 600

Qy 601 LELKNEQOTLRADMPFPESKQDYBESSWDSLSLCEVTSQKDVCLPKAAHQKEIDKING 660
 Db 601 LELKNEQOTLRADMPFPESKQDYBESSWDSLSLCEVTSQKDVCLPKAAHQKEIDKING 660

Qy 661 LEGSPVKDGLLKANGCMKVISIPTKALELMDQMTFKAEPGKPSAFEPATEMOKSVPNKAL 720
 Db 661 LEGSPVKDGLLKANGCMKVISIPTKALELMDQMTFKAEPGKPSAFEPATEMOKSVPNKAL 720

Qy 721 ELKNEQOTLRADMPFPESKQDYBESSWDSLSLCEVTSQKDVCLPKATHQKEIDKINGKL 780
 Db 721 ELKNEQOTLRADMPFPESKQDYBESSWDSLSLCEVTSQKDVCLPKATHQKEIDKINGKL 780

Qy 781 BESPDNDGFLKAPCRMKVSIPTKALELMDQMTFKAEPGKPSAFEPATEMOKSVPNKALE 840
 Db 781 BESPDNDGFLKAPCRMKVSIPTKALELMDQMTFKAEPGKPSAFEPATEMOKSVPNKALE 840

Qy 841 LKNEQOTLRADMPFPESKQDYBESSWDSLSLCEVTSQKDVCLPKATHQKEIDKINGKL 900
 Db 841 LKNEQOTLRADMPFPESKQDYBESSWDSLSLCEVTSQKDVCLPKATHQKEIDKINGKL 900

Qy 901 DSTLSLSKILTDTVHSCERARELQKDHCEQRTCKMEQMKKFCVLLKKLSEAKEIKSOLN 960
 Db 901 DSTLSLSKILTDTVHSCERARELQKDHCEQRTCKMEQMKKFCVLLKKLSEAKEIKSOLN 960

Qy 961 KVKWEQELCSVRLTLNQEEKRRNADILNEXIREELGRIEEQHRKELEVKQOLEQALRIQ 1020
 Db 961 KVKWEQELCSVRLTLNQEEKRRNADILNEXIREELGRIEEQHRKELEVKQOLEQALRIQ 1020

Qy 1021 DIEKLSVESNLNOYSHTHENENYLLHENCMLKKEIAMLKLEIATLKHOYQEKENKVFEDI 1080
 Db 1021 DIEKLSVESNLNOYSHTHENENYLLHENCMLKKEIAMLKLEIATLKHOYQEKENKVFEDI 1080

Qy 1081 KILKEKNAELQMTLKLKEESLTKEASQVSGQLKVLIAENTMTLSKLEKQDKKEILEASIE 1140
 Db 1081 KILKEKNAELQMTLKLKEESLTKEASQVSGQLKVLIAENTMTLSKLEKQDKKEILEASIE 1140

Qy 1141 SHHPRLASAVODHDQIVTSRKSQBPAPHIAGDACLRQKQNVDSSTIYNNNEVLHQPSEA 1200
 Db 1141 SHHPRLASAVODHDQIVTSRKSQBPAPHIAGDACLRQKQNVDSSTIYNNNEVLHQPSEA 1200

Qy 1201 ORKSKSLKINLYAGDALRENTLVSEHAQRDQRTQCMKEAHMYQNEODNVNKHSTQO 1260
 Db 1201 ORKSKSLKINLYAGDALRENTLVSEHAQRDQRTQCMKEAHMYQNEODNVNKHSTQO 1260

Qy 1261 ESLDQKFLQSQKNMVLQQLVHAHKKADNKSKITIDHFLERKQKHLLKKEKEEIEFNY 1320
 Db 1261 ESLDQKFLQSQKNMVLQQLVHAHKKADNKSKITIDHFLERKQKHLLKKEKEEIEFNY 1320

Qy 1321 NNHLKNRIYQVEKEKAETENS 1341
 Db 1321 NNHLKNRIYQVEKEKAETENS 1341

RESULT 5
 ADL93227
 ID ADL93227 standard; protein; 1341 AA.
 XX AC ADL93227;
 XX DT 20-MAY-2004 (first entry)
 XX DE Human breast cancer-associated polypeptide #49.
 XX KW gene therapy; protein therapy; vaccine; breast cancer; cancer; human.
 XX OS Homo sapiens.
 XX PN US2003166022-A1.
 XX PD 04-SEP-2003.
 XX PF 15-APR-2002;-2002US-00124805.
 XX PR 28-DEC-1998; 98US-00222575.
 XX PR 02-APR-1999; 99US-00285480.
 XX PR 23-JUN-1999; 99US-00339338.
 XX PR 02-SEP-1999; 99US-00389681.
 XX PR 03-NOV-1999; 99US-00433826.
 XX PR 17-APR-2000; 2000US-00551621.
 XX PR 08-JUN-2000; 2000US-00590751.

PR	22-JUN-2000; 2000US-00604287.
PR	20-JUL-2000; 2000US-00620405.
PR	13-APR-2001; 2001US-00834759.
PR	07-DEC-2001; 2001US-00007805.
PR	13-FEB-2002; 2002US-00076622.
XX	(CORI-) CORIXA CORP.
XX	
XX	Houghton RL, Sleath PR, Persing DH;
XX	
XX	WPI; 2003-874918/81.
DR	N-PSDB; ADL93226.
XX	
XX	An isolated oncogenic polypeptide useful for preventing, diagnosing and
PT	treating breast cancer.
PT	
XX	
XX	Example 12; SEQ ID NO 565; 294pp; English.
PS	
XX	The invention relates to an isolated breast cancer-associated
CC	polypeptide. The polypeptide may be used for the diagnosis and treatment
CC	of breast cancers. The methods are useful for detecting the presence of a
CC	cancer in a patient and treating a cancer in a patient. The present
CC	sequence represents the amino acid sequence of a human breast cancer-
CC	associated polypeptide.
XX	
XX	Sequence 1341 AA;
SQ	
	Query Match 100.0%; Score 6920; DB 7; Length 1341;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 1341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 MTKRKKTINLMIQDAQKRTALHWACVNGHEEVTFLVDRKCOLDVLGGEHRTPLMKALQC 60
Db	1 MTKRKKTINLMIQDAQKRTALHWACVNGHEEVTFLVDRKCOLDVLGGEHRTPLMKALQC 60
Qy	61 HOEACANLIDSGADINLVDVYGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 120
Db	61 HOEACANLIDSGADINLVDVYGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 120
Qy	121 LLSITKRSEQIVFLLIKNANANA VNKYKCTALMLAVCHGSGSEIVGMLLQONVDVFAADI 180
Db	121 LLSITKRSEQIVFLLIKNANANA VNKYKCTALMLAVCHGSGSEIVGMLLQONVDVFAADI 180
Qy	181 CGVTAEHYAVTCGFHHIHEQIMEYIRKLSKNHNTNPGTSGAGTDEAAPLAERTPDTAEE 240
Db	181 CGVTAEHYAVTCGFHHIHEQIMEYIRKLSKNHNTNPGTSGAGTDEAAPLAERTPDTAEE 240
Qy	241 SLVEKTPDEAAPLVERTPDTAESLVEKTPDEAASLVEGTSDKIQCLEKATSGKFFQSAAE 300
Db	241 SLVEKTPDEAAPLVERTPDTAESLVEKTPDEAASLVEGTSDKIQCLEKATSGKFFQSAAE 300
Qy	301 TPREITSPAKETSEKFTWPAKGRPKIAWEKKEDTPREIMSPAKETSEKFTWAAKGRPRK 360
Db	301 TPREITSPAKETSEKFTWPAKGRPKIAWEKKEDTPREIMSPAKETSEKFTWAAKGRPRK 360
Qy	361 IAWKKETPVKTCGVARVTSNTKVLKGRSKNIACTPKESSTKASANDORPFSSEKQBE 420
Db	361 IAWKKETPVKTCGVARVTSNTKVLKGRSKNIACTPKESSTKASANDORPFSSEKQBE 420
Qy	421 DEEYSCDSRLSFESSAKIQVCIPESITYQKWEINREVEPPPKPSAFKPAIEMQNSVPNK 480
Db	421 DEEYSCDSRLSFESSAKIQVCIPESITYQKWEINREVEPPPKPSAFKPAIEMQNSVPNK 480
Qy	481 AFELKNEQTLRADPMPFPSPSKQXDYBENSWDSESLCETYSQKDVCCLPKATHOKEIDKING 540
Db	481 AFELKNEQTLRADPMPFPSPSKQXDYBENSWDSESLCETYSQKDVCCLPKATHOKEIDKING 540
Qy	541 KLEESPNKQGLLKATCGMKVSIPTKALELKDMDTFAEPGPKPSAFEPATEMQKSPVNKA 600
Db	541 KLEESPNKQGLLKATCGMKVSIPTKALELKDMDTFAEPGPKPSAFEPATEMQKSPVNKA 600
Qy	601 LELKNEQTWRADIELPSESKQXDYBENSWDSESLCETVTSQKDVCCLPKAAHOKREIDKING 660

PR 02-APR-1999; 99US-00285480.
 PR 23-JUN-1999; 99US-00339338.
 PR 02-SEP-1999; 99US-00389681.
 PR 03-NOV-1999; 99US-00433826.
 PR 17-APR-2000; 2000US-00551621.
 PR 08-JUN-2000; 2000US-00590751.
 PR 22-JUN-2000; 2000US-00604287.
 PR 20-JUL-2000; 2000US-00620405.
 PR 13-APR-2001; 2001US-00834759.
 PR 07-DEC-2001; 2001US-00007805.
 PR 13-FEB-2002; 2002US-00076622.
 XX (CORI-) CORIXA CORP.
 XX
 XX Houghton RL, Sleath PR, Persing DH;
 XX
 DR WPI; 2003-874918/81.
 DR N-PSDB; ADL93228.
 XX
 PT An isolated oncogenic polypeptide useful for preventing, diagnosing and
 PT treating breast cancer.
 XX
 PS Example 12; SEQ ID NO 573; 294pp; English.
 XX
 CC The invention relates to an isolated breast cancer-associated
 CC polypeptide. The polypeptide may be used for the diagnosis and treatment
 CC of breast cancers. The methods are useful for detecting the presence of a
 CC cancer in a patient and treating a cancer in a patient. The present
 CC sequence represents the amino acid sequence of a human breast cancer-
 CC associated polypeptide.
 XX
 SQ Sequence 1349 AA;

Query Match 99.6%; Score 6892; DB 7; Length 1349;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1336; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TKRKTININLQDAQKRTALHAWCVNGHEEVVTVFLVDRKQCLDVLDCGHRTPLMKALQCH 61
 DB 10 TKRKTININLQDAQKRTALHAWCVNGHEEVVTVFLVDRKQCPDVLDCGHRTPLMKALQCH 69
 QY 62 QEACANILIDSGADINLVDVYGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPLL 121
 DB 70 QEACANILIDSGADINLVDVYGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPLL 129
 QY 122 LSITKRSQIVVEFLIKNANANVKNYKCTALMLAVCHGSSEIVGMLLQQNVDFVFAADIC 181
 DB 130 LSITKRSQIVVEFLIKNANANVKNYKCTALMLAVCHGLSEIVGMLLQQNVDFVFAADIC 189
 QY 182 GVTAEHYAVTCGFHHIHEQIMEYIRKLSKNHNTNPEGTSAGTPDEAAPLAERTPDTAES 241
 DB 190 GVTAEHYAVTCGFHHIHEQIMEYIRKLSKNHNTNPEGTSAGTPDEAAPLAERTPDTAES 249
 QY 242 LVEKTPDEAAPLVERTPTASLSVEKTPDEAASLVEGSDIKIQCLEKATSKGFSQABET 301
 DB 250 LVEKTPDEAAPLVERTPTASLSVEKTPDEAASLVEGSDIKIQCLEKATSKGFSQABET 309
 QY 302 PREITSPAKETSEKFTWPAKGRPKIAWEKKEDTTPREIMSPAKETSEKFTWAAKGRPKI 361
 DB 310 PREITSPAKETSEKFTWPAKGRPKIAWEKKEDTTPREIMSPAKETSEKFTWAAKGRPKI 369
 QY 362 AWEKKEPVKTCGVARVTSNKTIVLEKGRSKMIACPTKESSTKASANDQRPFSKQBED 421
 DB 370 AWEKKEPVKTCGVARVTSNKTIVLEKGRSKMIACPTKESSTKASANDQRPFSKQBED 429
 QY 422 BEYSCDSRLPSSAKIQVCIPESITYQKMEINREVEPPPKPSAPKPAIEMQNSVPNKA 481
 DB 430 BEYSCDSRLPSSAKIQVCIPESITYQKMEINREVEPPPKPSAPKPAIEMQNSVPNKA 489
 QY 482 FELKNEOTLRADMPFPSPKOKYEENSWSDESLSCTVSKQDVCLPKATHOKEIDKINGK 541
 DB 490 FELKNEOTLRADMPFPSPKOKYEENSWSDESLSCTVSKQDVCLPKATHOKEIDKINGK 549

QY 542 LEESPNDGGLLKATCGMKVSIPTKALELKMOTFKABPPGKPSAFEPATEMOKSVPNKAL 601
 DB 550 LEESPNDGGLLKATCGMKVSIPTKALELKMOTFKABPPGKPSAFEPATEMOKSVPNKAL 609
 QY 602 ELKNEQTRADEILPSESOKDYEENSWSDESLSCTVSKQDVCLPKAAHQKEIDKINGKL 661
 DB 610 ELKNEQTRADEILPSESOKDYEENSWSDESLSCTVSKQDVCLPKAAHQKEIDKINGKL 669
 QY 662 EGSPVKDGLLKANCCKMKVSIPTKALELMDMOTFKABPPKPSAFEPATEMOKSVPNKALE 721
 DB 670 EGSPVKDGLLKANCCKMKVSIPTKALELMDMOTFKABPPKPSAFEPATEMOKSVPNKALE 729
 QY 722 LKNEQTLRADEILPSESOKDYEENSWSDESLSCTVSKQDVCLPKATHOKEIDKINGKLE 781
 DB 730 LKNEQTLRADEILPSESOKDYEENSWSDESLSCTVSKQDVCLPKATHOKEIDKINGKLE 789
 QY 782 ESPNDGFLKAPCKMKVSIPTKALELMDMOTFKABPPKPSAFEPATEMOKSVPNKALEL 841
 DB 790 ESPNDGFLKAPCKMKVSIPTKALELMDMOTFKABPPKPSAFEPATEMOKSVPNKALEL 849
 QY 842 KNEQTLRADQMPFPSPKOKYEENSWSDESLSCTVSKQDVCPKATHOKEIDKINGKLE 901
 DB 850 KNEQTLRADQMPFPSPKOKYEENSWSDESLSCTVSKQDVCPKATHOKEIDKINGKLE 909
 QY 902 STLSKILDTVHSCERARELOKHCEORTGHMEQMKKFCVLKKLSEAKIKSOLENOK 961
 DB 910 STLSKILDTVHSCERARELOKHCEORTGHMEQMKKFCVLKKLSEAKIKSOLENOK 969
 QY 962 VKWEQELCSVRLTUNQBEKKRNADIINEKIRELSGRIEEOHKELEVKQOLEALRIQD 1021
 DB 970 VKWEQELCSVRLTUNQBEKKRNADIINEKIRELSGRIEEOHKELEVKQOLEALRIQD 1029
 QY 1022 IELKSVESNLNQVSHTHENENYLLHENCMLKKEIATLKLKLEIATLKHQYQEKENYFEDIK 1081
 DB 1030 IELKSVESNLNQVSHTHENENYLLHENCMLKKEIATLKLKLEIATLKHQYQEKENYFEDIK 1089
 QY 1082 ILKEKNAELQMTLKLKESLTKRASQYSGQLKVLIAENTMLTSKLKEKODKEILEAEIES 1141
 DB 1090 ILKEKNAELQMTLKLKESLTKRASQYSGQLKVLIAENTMLTSKLKEKODKEILEAEIES 1149
 QY 1142 HHPLASAVODHDOIVTSRKSQEPAFHTAGDACLQKMNVDVSTIYNNVHLHOPLSBAQ 1201
 DB 1150 HHPLASAVODHDOIVTSRKSQEPAFHTAGDACLQKMNVDVSTIYNNVHLHOPLSBAQ 1209
 QY 1202 RKSKSLINLNYAGDALRENTLVSEHAQDQRETCCQKKEAHEMYQNEQDNNVKNHTEQOE 1261
 DB 1210 RKSKSLINLNYAGDALRENTLVSEHAQDQRETCCQKKEAHEMYQNEQDNNVKNHTEQOE 1269
 QY 1262 SLQDKLFQLOSKNNWLOQLVHAHKADNKSKITIDHFLERKQHLLKEKNEIFNYN 1321
 DB 1270 SLQDKLFQLOSKNNWLOQLVHAHKADNKSKITIDHFLERKQHLLKEKNEIFNYN 1329
 QY 1322 NHLKNRIYQYEKEAETE 1339
 DB 1330 NHLKNRIYQYEKEAETE 1347
 RESULT 8
 AAU33351
 ID AAU33351 standard; protein; 1002 AA.
 XX
 AC AAU33351;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human breast cancer protein B726P from alternatively spliced cDNA.
 XX
 KW Human; ss; breast cancer protein; tumour; cancer; cytostatic;
 KW gene therapy; immunogen.
 XX
 OS Homo sapiens.
 XX
 PN WO200179286-A2.

XX 25-OCT-2001.
 XX 12-APR-2001; 2001WO-US012164.
 XX 17-APR-2000; 2000US-00551621.
 XX 08-JUN-2000; 2000US-00390751.
 XX 22-JUN-2000; 2000US-00604287.
 XX 20-JUL-2000; 2000US-00620405.
 XX (CORI-) CORIXA CORP.
 XX
 XX Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
 XX WPI; 2001-611721/70.
 XX N-PSDB; AAS47411.
 XX
 XX Breast Tumor Proteins and nucleic acids useful for the prevention,
 XX diagnosis and treatment of breast cancer.
 XX
 XX Claim 3; Page 281-283; 297pp; English.
 XX
 XX The invention relates to isolated breast tumour proteins and nucleic
 XX acids that encode them, including immunogenic fragments of the proteins.
 XX Also included are expression vectors expressing the proteins, transformed
 XX cells and antibodies raised against the proteins or an antigen presenting
 XX cell expressing the protein. The proteins and nucleic acids may be used
 XX in the prevention, diagnosis and treatment of diseases associated with
 XX inappropriate breast tumour protein expression, i.e. breast tumours and
 XX breast cancer e.g. by gene therapy. The nucleic acids and their
 XX complements may also be used as DNA probes in diagnostic assays to detect
 XX and quantitate the presence of similar nucleic acids in samples, and
 XX therefore which patients may be in need of restorative therapy. The
 XX proteins, nucleic acids and antibodies may be used in assays to identify
 XX modulators (e.g. antagonists) of breast tumour protein expression and
 XX activity. The antibodies and antagonists may also be used to down
 XX regulate expression and activity. The antibodies may also be used as
 XX diagnostic agents for detecting the presence of the proteins in samples
 XX (e.g. by enzyme linked immunosorbent assay (ELISA)) and in other immuno-
 XX purification diagnostic techniques. The present sequence is a breast
 XX tumour protein encoded by a cDNA from a breast tumour cDNA library
 XX isolated by subtractive hybridisation against a normal breast cDNA
 XX library
 XX
 XX SQ Sequence 1002 AA;
 Query Match 74.1%; Score 5130; DB 4; Length 1002;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 Qy 340 MSPAKETSEKFTWAAKGRPKIAEKKETPVKTGCVARVTSNKTVKLEKGRSKMIACPTK 399
 Db 1 MSPAKETSEKFTWAAKGRPKIAEKKETPVKTGCVARVTSNKTVKLEKGRSKMIACPTK 60
 Qy 400 ESSTKASANDORPSESQDEDEYSCDSRSLFESSAKIQVICIPESIQKWEINREVEE 459
 Db 61 ESSTKASANDORPSESQDEDEYSCDSRSLFESSAKIQVICIPESIQKWEINREVEE 120
 Qy 460 PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADMPPEPSKQKDYBENSWSLCEV 519
 Db 121 PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADMPPEPSKQKDYBENSWSLCEV 180
 Qy 520 SQKDVCLPKATHQKEIDKINGKLEBSPNKGILLKATCGMKVSIPTKALELMDMTFFKAEP 579
 Db 181 SQKDVCLPKATHQKEIDKINGKLEBSPNKGILLKATCGMKVSIPTKALELMDMTFFKAEP 240
 Qy 580 PGKPSAFEPATEMOKSVPNKALELKNQETWADAILPSESQKDYBENSWSLCEV 639
 Db 241 PGKPSAFEPATEMOKSVPNKALELKNQETWADAILPSESQKDYBENSWSLCEV 300
 Qy 640 QKDVCLPKAAHQKEIDKINGKLEBSPNKGILLKATCGMKVSIPTKALELMDMTFFKAEP 699
 Db 301 QKDVCLPKAAHQKEIDKINGKLEBSPNKGILLKATCGMKVSIPTKALELMDMTFFKAEP 360

Qy 700 EKPSAFEPATEMOKSVPNKALELKNQETWADAILPSESQKDYBENSWSLCEV 759
 Db 361 EKPSAFEPATEMOKSVPNKALELKNQETWADAILPSESQKDYBENSWSLCEV 420
 Qy 760 KDVCLPKATHQKEIDKINGKLEBSPNKGILLKATCGMKVSIPTKALELMDMTFFKAEP 819
 Db 421 KDVCLPKATHQKEIDKINGKLEBSPNKGILLKATCGMKVSIPTKALELMDMTFFKAEP 480
 Qy 820 KPSAFEPATEMOKSVPNKALELKNQETWADAILPSESQKDYBENSWSLCEV 879
 Db 481 KPSAFEPATEMOKSVPNKALELKNQETWADAILPSESQKDYBENSWSLCEV 540
 Qy 880 DVCVPKATHQKEMDKISGKLESDSTLSKILDTVHSCERARELQKDHCEQRTGKMEQMKKK 939
 Db 541 DVCVPKATHQKEMDKISGKLESDSTLSKILDTVHSCERARELQKDHCEQRTGKMEQMKKK 600
 Qy 940 FCVLKCKLSEAKEIKSQLENQKQVWEQELCSVRLTLNQEBEKKRRNADILNEKIREELGRI 999
 Db 601 FCVLKCKLSEAKEIKSQLENQKQVWEQELCSVRLTLNQEBEKKRRNADILNEKIREELGRI 660
 Qy 1000 BEQHRKELEVQKQLEQALRIQDIELKSVESNLNOVSHTHENENYLLHENCMLKKEIAMLK 1059
 Db 661 BEQHRKELEVQKQLEQALRIQDIELKSVESNLNOVSHTHENENYLLHENCMLKKEIAMLK 720
 Qy 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 1119
 Db 721 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 780
 Qy 1120 TMLTSKLKEKQDKLEIAEIESHPRLASAVQDHDQIVTSRKSPAPAFIAGDACLQRM 1179
 Db 781 TMLTSKLKEKQDKLEIAEIESHPRLASAVQDHDQIVTSRKSPAPAFIAGDACLQRM 840
 Qy 1180 NVDVSTIYNNNEVLHQPULSEAQRSKSLKINLTAGDALRENTLVSEHAQRDQRETQCOM 1239
 Db 841 NVDVSTIYNNNEVLHQPULSEAQRSKSLKINLTAGDALRENTLVSEHAQRDQRETQCOM 900
 Qy 1240 KEAEHMYQNEODNVNKHTEQOESLDQKLFQLSKNMMLQQQLVHAHKKADNKSKITIDH 1299
 Db 901 KEAEHMYQNEODNVNKHTEQOESLDQKLFQLSKNMMLQQQLVHAHKKADNKSKITIDH 960
 Qy 1300 FLERKMQHLLKKEKNEEIFNYYNNHLKNRIYQYEKEKATENS 1341
 Db 961 FLERKMQHLLKKEKNEEIFNYYNNHLKNRIYQYEKEKATENS 1002
 RESULT 9
 ABG78918
 ID ABG78918 standard; protein; 1002 AA.
 XX AC ABG78918;
 XX XX
 XX DT 15-NOV-2002 (first entry)
 XX DE Human breast tumour polypeptide #10.
 XX XX Human; breast tumour protein; breast cancer; cytostatic; vaccine.
 XX KW Homo sapiens.
 XX OS
 XX XX US2002085998-A1.
 XX XX
 XX PD 04-JUL-2002.
 XX PF 13-APR-2001; 2001US-00834759.
 XX XX
 XX PR 28-DEC-1998; 98US-00222575.
 XX PR 02-APR-1999; 99US-00285480.
 XX PR 23-JUN-1999; 99US-00339338.
 XX PR 02-SEP-1999; 99US-00389681.
 XX PR 03-NOV-1999; 99US-00433826.
 XX PR 17-APR-2000; 2000US-00551621.
 XX PR 08-JUN-2000; 2000US-00590751.

CC cancer). The method and sequences of the invention are useful for
CC stimulating and/or expanding T cells specific for a tumour protein,
CC detecting the presence of cancer, stimulating an immune response in a
CC patient and treating breast cancer. The present amino acid sequence
CC represents a human tumour-related protein
XX
XX Sequence 1002 AA;

Query Match 74.1%; Score 5130; DB 6; Length 1002;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 340 MSPAKETSEKFTWAAGPRKIAWEKKEPTVKTCGVARVTSNKTIVLEKGRSKMIACPTK 399
DB 1 MSPAKETSEKFTWAAGPRKIAWEKKEPTVKTCGVARVTSNKTIVLEKGRSKMIACPTK 60

QY 400 ESSTKASANDORFPSESKQEDDEEYSCDSRSLFPSSAKIQVCIPESIIYQKWEINREVEE 459
DB 61 ESSTKASANDORFPSESKQEDDEEYSCDSRSLFPSSAKIQVCIPESIIYQKWEINREVEE 120

QY 460 PPKPSAFKPAIEMQNSVPNKAFLKNEQTLRADPMFPPEKQKDYENSWSSESCLCTV 519
DB 121 PPKPSAFKPAIEMQNSVPNKAFLKNEQTLRADPMFPPEKQKDYENSWSSESCLCTV 180

QY 520 SKQVCLPKATHQKEIDKINGKLEBSPNKGDLKATCGMKVSIPTKALELMDQTFKAEP 579
DB 181 SKQVCLPKATHQKEIDKINGKLEBSPNKGDLKATCGMKVSIPTKALELMDQTFKAEP 240

QY 580 PGKSAFEPATEMOKSVNKALELKNQETWRADELTPSESKQKDYENSWSSESCLCTV 639
DB 241 PGKSAFEPATEMOKSVNKALELKNQETWRADELTPSESKQKDYENSWSSESCLCTV 300

QY 640 QKDVCPLKAAHQKEIDKINGKLEBSPNKGDLKATCGMKVSIPTKALELMDQTFKAEP 699
DB 301 QKDVCPLKAAHQKEIDKINGKLEBSPNKGDLKATCGMKVSIPTKALELMDQTFKAEP 360

QY 700 EKPSAFEPATEMOKSVNKALELKNQETWRADELTPSESKQKDYENSWSSESCLCTV 759
DB 361 EKPSAFEPATEMOKSVNKALELKNQETWRADELTPSESKQKDYENSWSSESCLCTV 420

QY 760 KDVCPLPKATHQKEIDKINGKLEBSPNKGDLKATCGMKVSIPTKALELMDQTFKAEP 819
DB 421 KDVCPLPKATHQKEIDKINGKLEBSPNKGDLKATCGMKVSIPTKALELMDQTFKAEP 480

QY 820 KPSAFEPATEMOKSVNKALELKNQETWRADELTPSESKQKDYENSWSSESCLCTV 879
DB 481 KPSAFEPATEMOKSVNKALELKNQETWRADELTPSESKQKDYENSWSSESCLCTV 540

QY 880 DVCVPKATHQKEMDKISGKLEBSPNKGDLKATCGMKVSIPTKALELMDQTFKAEP 939
DB 541 DVCVPKATHQKEMDKISGKLEBSPNKGDLKATCGMKVSIPTKALELMDQTFKAEP 600

QY 940 FCVLKKLSEAKEIKSLENOKVWEQLCSVRLTNOEERKNADILNEKIREELGRI 999
DB 601 FCVLKKLSEAKEIKSLENOKVWEQLCSVRLTNOEERKNADILNEKIREELGRI 660

QY 1000 EEOHRKELEVKQLEQALRIODIELKSVESNLNOVSHTHENYLLHNCMLKKEIAMLK 1059
DB 661 EEOHRKELEVKQLEQALRIODIELKSVESNLNOVSHTHENYLLHNCMLKKEIAMLK 720

QY 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 1119
DB 721 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 780

QY 1120 TMLTSKLEKQDKEILEAEISHHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQRM 1179
DB 781 TMLTSKLEKQDKEILEAEISHHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQRM 840

QY 1180 NVDVSSITYNNEVLHQPSEAKRSKSLKINLYAGDALRENTLVSEHAQRDQRETQCM 1239
DB 841 NVDVSSITYNNEVLHQPSEAKRSKSLKINLYAGDALRENTLVSEHAQRDQRETQCM 900

QY 1240 KEAEHMYQNEQDNVKNKHTEQOESLDOKLFOLQSKNMWLQQLVHAHKADNKSKITIDIH 1299

DB 901 KEAEHMYQNEQDNVKNKHTEQOESLDOKLFOLQSKNMWLQQLVHAHKADNKSKITIDIH 960
QY 1300 FLERKMQHLLKEKNEEIFFNNHLLKNRIYQVEKEKAETENS 1341
DB 961 FLERKMQHLLKEKNEEIFFNNHLLKNRIYQVEKEKAETENS 1002

RESULT 11
ADL93137
ID ADL93137 standard; protein; 1002 AA.
XX
AC ADL93137;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human breast cancer-associated polypeptide #10.
XX
KW gene therapy; protein therapy; vaccine; breast cancer; cancer; human.
XX
OS Homo sapiens.
XX
PN US2003166022-A1.
XX
PD 04-SEP-2003.
XX
PF 15-APR-2002; 2002US-00124805.
XX
PR 28-DEC-1998; 98US-00222575.
PR 02-APR-1999; 99US-00285480.
PR 23-JUN-1999; 99US-00339338.
PR 02-SEP-1999; 99US-00389681.
PR 03-NOV-1999; 99US-00433826.
PR 17-APR-2000; 2000US-00551621.
PR 08-JUN-2000; 2000US-00590751.
PR 22-JUN-2000; 2000US-00604287.
PR 20-JUL-2000; 2000US-00620405.
PR 13-APR-2001; 2001US-00834759.
PR 07-DEC-2001; 2001US-00007805.
PR 13-FEB-2002; 2002US-00076622.
XX
PA (CORI-) CORIXA CORP.
XX
PI Houghton RL, Sleath PR, Persing DH;
XX
XX WPI: 2003-874918/81.
DR N-PSDB; ADL93136.
XX
PT An isolated oncogenic polypeptide useful for preventing, diagnosing and
XX treating breast cancer.
XX
XX Example 1; SEQ ID NO 475; 294pp; English.
XX
XX The invention relates to an isolated breast cancer-associated
XX polypeptide. The polypeptide may be used for the diagnosis and treatment
XX of breast cancers. The methods are useful for detecting the presence of a
XX cancer in a patient and treating a cancer in a patient. The present
XX sequence represents the amino acid sequence of a human breast cancer-
XX associated polypeptide.
XX
SQ Sequence 1002 AA;

Query Match 74.1%; Score 5130; DB 7; Length 1002;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 340 MSPAKETSEKFTWAAGPRKIAWEKKEPTVKTCGVARVTSNKTIVLEKGRSKMIACPTK 399
DB 1 MSPAKETSEKFTWAAGPRKIAWEKKEPTVKTCGVARVTSNKTIVLEKGRSKMIACPTK 60

QY 400 ESSTKASANDORFPSESKQEDDEEYSCDSRSLFPSSAKIQVCIPESIIYQKWEINREVEE 459
DB 61 ESSTKASANDORFPSESKQEDDEEYSCDSRSLFPSSAKIQVCIPESIIYQKWEINREVEE 120


```
QY 460 PPKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 519
Db 121 PPKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 180
QY 520 SOKDVCLPKATHQKEIDKINGKLEESPNKDGILLKATCGMKVSIPTKALELMDQTFKAEP 579
Db 181 SOKDVCLPKATHQKEIDKINGKLEESPNKDGILLKATCGMKVSIPTKALELMDQTFKAEP 240
QY 580 PGKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 639
Db 241 PGKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 300
QY 640 QKDVCLPKAAHOKETDKINGKLESPVKGILLKANGCMKVSIPTKALELMDQTFKAEP 699
Db 301 QKDVCLPKAAHOKETDKINGKLESPVKGILLKANGCMKVSIPTKALELMDQTFKAEP 360
QY 700 EKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 759
Db 361 EKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 420
QY 760 KDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMVSIPTKALELMDQTFKAEP 819
Db 421 KDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMVSIPTKALELMDQTFKAEP 480
QY 820 KPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 879
Db 481 KPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 540
QY 880 DVCVPKATHQKEMDKISGKLEDSLSKILDTVHSCERARELQKDHCEQRTGKMEQMKK 939
Db 541 DVCVPKATHQKEMDKISGKLEDSLSKILDTVHSCERARELQKDHCEQRTGKMEQMKK 600
QY 940 FCVLKKLSEAKEIKSLENQKWEQBELCSVRLTLNOEEKRRNADILNEKIREELGRI 999
Db 601 FCVLKKLSEAKEIKSLENQKWEQBELCSVRLTLNOEEKRRNADILNEKIREELGRI 660
QY 1000 BEQHKELFVKOOLEQALRIQDIELKVESNLNOVSHTHENENYLLHNCMLKKEIAMLK 1059
Db 661 BEQHKELFVKOOLEQALRIQDIELKVESNLNOVSHTHENENYLLHNCMLKKEIAMLK 720
QY 1060 LEIATLKHQYQKKNYPEDIKILKEKNAELQMTLKLKEESITKRASQYSGQLKVLIAEN 1119
Db 721 LEIATLKHQYQKKNYPEDIKILKEKNAELQMTLKLKEESITKRASQYSGQLKVLIAEN 780
QY 1120 TMLTSKLKEQKQKEILEABIESHHPRASAVODHQIVTSRKSQBPFAHIAGDACLQRM 1179
Db 781 TMLTSKLKEQKQKEILEABIESHHPRASAVODHQIVTSRKSQBPFAHIAGDACLQRM 840
QY 1180 NVDVSTIYNNVHLQPLSEAKRSKSLKINLYAGDALRENTLVSEHAQRDQRTQCM 1239
Db 841 NVDVSTIYNNVHLQPLSEAKRSKSLKINLYAGDALRENTLVSEHAQRDQRTQCM 900
QY 1240 KEAEHMYQNEQDNVANKHTEQESLQKLFQLQSKNMWLQQLVHAHKADNKSKITIDH 1299
Db 901 KEAEHMYQNEQDNVANKHTEQESLQKLFQLQSKNMWLQQLVHAHKADNKSKITIDH 960
QY 1300 FLERKQVHLLKKEKEEIPYNNHLKNRIYQYKEKAETENS 1341
Db 961 FLERKQVHLLKKEKEEIPYNNHLKNRIYQYKEKAETENS 1002

RESULT 12
ADE44427
ID ADE44427 standard; protein; 1002 AA.
XX
AC ADE44427;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human breast cancer protein #10.
XX
KW human; breast tumour; cancer; vaccine; T cell stimulator;
```

```
T cell expander.
XX
OS Homo sapiens.
XX
PN US2003104366-A1.
XX
PD 05-JUN-2003.
XX
PF 17-APR-2000; 2000US-00551621.
XX
PR 28-DEC-1998; 98US-00222575.
PR 02-APR-1999; 99US-00285480.
PR 23-JUN-1999; 99US-00339338.
PR 02-SEP-1999; 99US-00389681.
PR 03-NOV-1999; 99US-00433826.
XX
(JIAN/) JIANG Y.
PA (DILL/) DILLON D C.
PA (MITC/) MITCHAM J L.
PA (XUJJ/) XU J.
PA (HARL/) HARLOCKER S L.
XX
Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL;
WPI; 2004-020270/02.
N-PSDB; ADE44426.
Novel isolated polypeptide comprising immunogenic portion of breast tumor protein or its variant, useful for formulating vaccines for inhibiting cancer development in a patient.
Example 1; SEQ ID NO 475; 217pp; English.
The invention relates to an isolated polypeptide comprising at least an immunogenic portion of a breast tumour protein. The polynucleotide, its fusion protein or the polynucleotide encoding it, a vaccine comprising the fusion protein or the polynucleotide encoding it, an isolated T cell population comprising T cells specific for a breast tumour protein, and a method for removing tumour cells from a biological sample is useful for inhibiting the development of a cancer in a patient. The polypeptide is useful for stimulating and/or expanding T cells specific for a breast tumour protein. Stimulating and/or expanding T cells specific for a breast tumour protein is useful for inhibiting the development of a cancer in a patient. The method additionally involves the step of cloning at least one proliferated cell and then administering the cloned T cells to the patient. The present sequence represents a breast cancer protein.
Sequence 1002 AA;
Query Match 74.1%; Score 5130; DB 8; Length 1002;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 340 MSPAKETSEKPTWAAKGRPRKIAWEKKEPTVKTCVARVTSNKTIVLEKSGSKMIACPTK 399
Db 1 MSPAKETSEKPTWAAKGRPRKIAWEKKEPTVKTCVARVTSNKTIVLEKSGSKMIACPTK 60
QY 400 ESSTKASANDQRPFSSEKQEEDEYSCDSRSLFPSSAKIOVCIPESIIYQKWEINREVEE 459
Db 61 ESSTKASANDQRPFSSEKQEEDEYSCDSRSLFPSSAKIOVCIPESIIYQKWEINREVEE 120
QY 460 PPKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 519
Db 121 PPKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 180
QY 520 SOKDVCLPKATHQKEIDKINGKLEESPNKDGILLKATCGMKVSIPTKALELMDQTFKAEP 579
Db 181 SOKDVCLPKATHQKEIDKINGKLEESPNKDGILLKATCGMKVSIPTKALELMDQTFKAEP 240
QY 580 PGKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 639
Db 241 PGKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 300
```

```
Qy 640 QKDVCLPKAAHQKEIDKINGKLESPVKDGLLKANCCKMVKSIPTKALELMDMOTFKAEPP 699
Db 301 QKDVCLPKAAHQKEIDKINGKLESPVKDGLLKANCCKMVKSIPTKALELMDMOTFKAEPP 360
Qy 700 EKPSAFPAIEMQKSVNPKALELNEQTLRADEILPSESKQKDYESSWDSSESLCETVSQ 759
Db 361 EKPSAFPAIEMQKSVNPKALELNEQTLRADEILPSESKQKDYESSWDSSESLCETVSQ 420
Qy 760 KDVCLPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPP 819
Db 421 KDVCLPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPP 480
Qy 820 KPSAFPAIEMQKSVNPKALELNEQTLRADEILPSESKQKDYESSWDSSESLCETVSQ 879
Db 481 KPSAFPAIEMQKSVNPKALELNEQTLRADEILPSESKQKDYESSWDSSESLCETVSQ 540
Qy 880 DVCVPKATHQEMKISKLEDSLSKILDTVHSCERARELOKDHCEQRTGKWEQMKKK 939
Db 541 DVCVPKATHQEMKISKLEDSLSKILDTVHSCERARELOKDHCEQRTGKWEQMKKK 600
Qy 940 FCVLKKLSEAKEIKSQLENOVKWQELCSVRLTLNOEESKRRNADILNEKIREELGRI 999
Db 601 FCVLKKLSEAKEIKSQLENOVKWQELCSVRLTLNOEESKRRNADILNEKIREELGRI 660
Qy 1000 EEOHKELEVKQLEQALRIQDIELKSVESNLNOVSHTHENYLLHENCMLKKEIAMLK 1059
Db 661 EEOHKELEVKQLEQALRIQDIELKSVESNLNOVSHTHENYLLHENCMLKKEIAMLK 720
Qy 1060 LEIATLKHQYQEKKNYPEDIKILKEKNAELQMTLKLKEESLTRKASQYSGQLKVLIAEN 1119
Db 721 LEIATLKHQYQEKKNYPEDIKILKEKNAELQMTLKLKEESLTRKASQYSGQLKVLIAEN 780
Qy 1120 TMLTSLKEKQDKETLEAEIESHPRLASAVQDDHDIIVTSRKSQEPAPHIAAGDACLQRM 1179
Db 781 TMLTSLKEKQDKETLEAEIESHPRLASAVQDDHDIIVTSRKSQEPAPHIAAGDACLQRM 840
Qy 1180 NVDVSSITYNNEVLHQLPSEARQSKSLKINLVAGDALRENTLVSEHAQRDQRETQCM 1239
Db 841 NVDVSSITYNNEVLHQLPSEARQSKSLKINLVAGDALRENTLVSEHAQRDQRETQCM 900
Qy 1240 KEAHEMYQNEQNVNKHTEQESLDQKLFLQSKNMWLQOQLVHAHKADNKSKITIDIH 1299
Db 901 KEAHEMYQNEQNVNKHTEQESLDQKLFLQSKNMWLQOQLVHAHKADNKSKITIDIH 960
Qy 1300 FLERQWQHLLKEXNEEITFNYNHLLKNRIYQYEKEKAETENS 1341
Db 961 FLERQWQHLLKEXNEEITFNYNHLLKNRIYQYEKEKAETENS 1002

RESULT 13
AAU33357
* ID AAU33357 standard; protein; 1095 AA.
AC AAU33357;
XX
XX
XX 18-DEC-2001 (first entry)
XX
DE Human breast cancer protein B726P fusion protein #1.
XX
XX Human; ss; breast cancer protein; tumour; cancer; cytostatic;
XX gene therapy; immunogen.
XX
XX Homo sapiens.
XX
XX WO200179286-A2.
XX
XX 25-OCT-2001.
XX
XX 12-APR-2001; 2001WO-US012164.
XX
XX 17-APR-2000; 2000US-00551621.
XX
XX 08-JUN-2000; 2000US-00590751.
```

```
PR 22-JUN-2000; 2000US-00604287.
PR 20-JUL-2000; 2000US-00620405.
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
XX WPI; 2001-611721/70.
DR N-PSDB; AAS47421.
XX
XX Breast Tumor Proteins and nucleic acids useful for the prevention,
PT diagnosis and treatment of breast cancer.
XX
XX Claim 22; Page 292-295; 297pp; English.
XX
XX The invention relates to isolated breast tumour proteins and nucleic
CC acids that encode them, including immunogenic fragments of the proteins.
CC Also included are expression vectors expressing the proteins, transformed
CC cells and antibodies raised against the proteins or an antigen presenting
CC cell expressing the protein. The proteins and nucleic acids may be used
CC in the prevention, diagnosis and treatment of diseases associated with
CC inappropriate breast tumour protein expression, i.e. breast tumours and
CC breast cancer e.g. by gene therapy. The nucleic acids and their
CC complements may also be used as DNA probes in diagnostic assays to detect
CC and quantitate the presence of similar nucleic acids in samples, and
CC therefore which patients may be in need of restorative therapy. The
CC proteins, nucleic acids and antibodies may be used in assays to identify
CC modulators (e.g. antagonists) of breast tumour protein expression and
CC activity. The antibodies and antagonists may also be used to down
CC regulate expression and activity. The antibodies may also be used as
CC diagnostic agents for detecting the presence of the proteins in samples
CC (e.g. by enzyme linked immunosorbant assay (ELISA)) and in other immuno-
CC purification diagnostic techniques. The present sequence is a breast
CC tumour protein encoded by a cDNA from a breast tumour cDNA library
CC isolated by subtractive hybridisation against a normal breast cDNA
CC library
XX
XX Sequence 1095 AA;
SQ
Query Match 74.1%; Score 5130; DB 4; Length 1095;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Qy 340 MSPAKETSEKFTWAAGRPRIAKWEKETPVKTGCVARVTSNKTVKLEKSGSKMIACPTK 399
Db 94 MSPAKETSEKFTWAAGRPRIAKWEKETPVKTGCVARVTSNKTVKLEKSGSKMIACPTK 153
Qy 400 ESSTKASANDQRPFPSESKQEEDEYSCDSRSLFPSSAKIQVCIPESIIYQKMEINREVEE 459
Db 154 ESSTKASANDQRPFPSESKQEEDEYSCDSRSLFPSSAKIQVCIPESIIYQKMEINREVEE 213
Qy 460 PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADPMFPFESKQKDYESSWDSSESLCETV 519
Db 214 PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADPMFPFESKQKDYESSWDSSESLCETV 273
Qy 520 SQKDVCLPKATHQKEIDKINGKLESPNKGDLKATCGMKVSIPTKALELMDMOTFKAEPP 579
Db 274 SQKDVCLPKATHQKEIDKINGKLESPNKGDLKATCGMKVSIPTKALELMDMOTFKAEPP 333
Qy 580 PGKPSAFEPATEMOKSVNPKALELNEQTLRADEILPSESKQKDYESSWDSSESLCETVS 639
Db 334 PGKPSAFEPATEMOKSVNPKALELNEQTLRADEILPSESKQKDYESSWDSSESLCETVS 393
Qy 640 QKDVCLPKAAHQKEIDKINGKLESPVKDGLLKANCCKMVKSIPTKALELMDMOTFKAEPP 699
Db 394 QKDVCLPKAAHQKEIDKINGKLESPVKDGLLKANCCKMVKSIPTKALELMDMOTFKAEPP 453
Qy 700 EKPSAFPAIEMQKSVNPKALELNEQTLRADEILPSESKQKDYESSWDSSESLCETVSQ 759
Db 454 EKPSAFPAIEMQKSVNPKALELNEQTLRADEILPSESKQKDYESSWDSSESLCETVSQ 513
Qy 760 KDVCLPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPP 819
Db 819 KDVCLPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPP 819
```

Db 514 KDVCPLPKAXHQEKIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMQTFKAEPPPE 573
Qy 820 KPSAFEPALMOKSVNPKALELKNQOTLRADQMPFSESKQKVRENSWDSLSLRETVSQK 879
Db 574 KPSAFEPALMOKSVNPKALELKNQOTLRADQMPFSESKQKVRENSWDSLSLRETVSQK 633
Qy 880 DVCVPKATHQEMDKISGKLESDTSLSKILDTVHSCERARELQKDHCEQRTGKMEQMKKK 939
Db 634 DVCVPKATHQEMDKISGKLESDTSLSKILDTVHSCERARELQKDHCEQRTGKMEQMKKK 693
Qy 940 FCVLKKLSEAKEIKSQLENQKVWEQELCSVRLTLNOEEKRRNADILNEKIREELGRI 999
Db 694 FCVLKKLSEAKEIKSQLENQKVWEQELCSVRLTLNOEEKRRNADILNEKIREELGRI 753
Qy 1000 EEOHREKLEVKQOOLQALRIODIELKSVESNLNQVSHTHENENYLLHNCMLKKEIAMLK 1059
Db 754 EEOHREKLEVKQOOLQALRIODIELKSVESNLNQVSHTHENENYLLHNCMLKKEIAMLK 813
Qy 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 1119
Db 814 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 873
Qy 1120 TMLTSKLKEKQDKKEILEAEIESHPRLASAVQDHDQIVTSRKSPAPFIAAGDAQLQRM 1179
Db 874 TMLTSKLKEKQDKKEILEAEIESHPRLASAVQDHDQIVTSRKSPAPFIAAGDAQLQRM 933
Qy 1180 NVDVSSSTIYNNEVLHQPILSEAKRSKSLKINLYAGDALRENTLVSEHAQRDQRTQCOM 1239
Db 934 NVDVSSSTIYNNEVLHQPILSEAKRSKSLKINLYAGDALRENTLVSEHAQRDQRTQCOM 993
Qy 1240 KEAEHMYQNEQDNVKNKHTEQEESLDQKLFQLOSKNMWLLQOOLVHAHKADNKSKITIDIH 1299
Db 994 KEAEHMYQNEQDNVKNKHTEQEESLDQKLFQLOSKNMWLLQOOLVHAHKADNKSKITIDIH 1053
Qy 1300 FLERQMOHLLKKEKNEEFNYYNHLKNRIYQVEKEKAETENS 1341
Db 1054 FLERQMOHLLKKEKNEEFNYYNHLKNRIYQVEKEKAETENS 1095

RESULT 14

ABG78924
ID ABG78924 standard; protein; 1095 AA.

XX AC ABG78924;

XX DT 15-NOV-2002 (first entry)

XX DE Human breast tumour polypeptide #15.

XX KW Human; breast tumour protein; breast cancer; cytostatic; vaccine.

XX OS Homo sapiens.

XX PN US2002085998-A1.

XX PD 04-JUL-2002.

XX PF 13-APR-2001; 2001US-00834759.

XX PR 28-DEC-1998; 98US-00222575.

XX PR 02-APR-1999; 99US-00285480.

XX PR 23-JUN-1999; 99US-00339338.

XX PR 02-SEP-1999; 99US-00389681.

XX PR 03-NOV-1999; 99US-00433826.

XX PR 17-APR-2000; 2000US-00551621.

XX PR 08-JUN-2000; 2000US-00590751.

XX PR 22-JUN-2000; 2000US-00604287.

XX PR 20-JUL-2000; 2000US-00620405.

XX PA (CORI-) CORIXA CORP.

XX PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
PI Henderson RA;

XX WPI; 2002-635657/68.
DR N-PSDB; ABS64022.
XX Novel breast cancer polynucleotides and polypeptides encoded by the
PT polynucleotides, useful for detecting the presence of breast cancer in a
PT patient, and in pharmaceutical compositions, for treating breast cancer.
XX Disclosure; Page 223-225; 247pp; English.
XX The invention relates to an isolated breast tumour polynucleotide and the
CC polypeptide it encodes. The polynucleotide and polypeptide are useful for
CC detecting the presence of breast cancer in a patient, and in
CC pharmaceutical compositions for treating breast cancer. The sequences are
CC useful for stimulating an immune response in a patient and can therefore
CC be used in production of vaccines. The sequences are also useful for
CC detecting the presence of a cancer in a patient, by obtaining a
CC biological sample from the patient, contacting the biological sample with
CC a composition of the invention and detecting the amount of polynucleotide
CC that hybridizes to the sample. This sequence represents a human breast
CC tumour polypeptide of the invention
XX SQ Sequence 1095 AA;

Query Match 74.1%; Score 5130; DB 5; Length 1095;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 340 MSPAKETSEKPTWAAGRPRIAEKKEKTPVTCGVARVTSNKTQVLEKSGSKMIACPTK 399

Db 94 MSPAKETSEKPTWAAGRPRIAEKKEKTPVTCGVARVTSNKTQVLEKSGSKMIACPTK 153

Qy 400 ESSSTKASANDORFSESKQDEEYSCDSRLSFESSAKIQVCIPESIIYQKWEINREVEE 459

Db 154 ESSSTKASANDORFSESKQDEEYSCDSRLSFESSAKIQVCIPESIIYQKWEINREVEE 213

Qy 460 PPKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADPMPFPESKQKDYENSWSLSELCETV 519

Db 214 PPKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADPMPFPESKQKDYENSWSLSELCETV 273

Qy 520 SQKDVCLPKATHQEKIDKINGKLEESPNKDGLLKATCGMKVSIPTKALELMDMQTFKAEPP 579

Db 274 SQKDVCLPKATHQEKIDKINGKLEESPNKDGLLKATCGMKVSIPTKALELMDMQTFKAEPP 333

Qy 580 PGKPSAFEPATEMOKSVNPKALELKNQOTLRADQMPFSESKQKVRENSWDSLSLRETVSQK 639

Db 334 PGKPSAFEPATEMOKSVNPKALELKNQOTLRADQMPFSESKQKVRENSWDSLSLRETVSQK 393

Qy 640 QKDVCLPKAAHQEKIDKINGKLEESPVKDGILLKANCGRMKVSIPTKALELMDMQTFKAEPP 699

Db 394 QKDVCLPKAAHQEKIDKINGKLEESPVKDGILLKANCGRMKVSIPTKALELMDMQTFKAEPP 453

Qy 700 EKPSAFEPALMOKSVNPKALELKNQOTLRADQMPFSESKQKVRENSWDSLSLRETVSQK 759

Db 454 EKPSAFEPALMOKSVNPKALELKNQOTLRADQMPFSESKQKVRENSWDSLSLRETVSQK 513

Qy 760 KDVCPLPKATHQEKIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMQTFKAEPP 819

Db 514 KDVCPLPKATHQEKIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMQTFKAEPP 573

Qy 820 KPSAFEPALMOKSVNPKALELKNQOTLRADQMPFSESKQKVRENSWDSLSLRETVSQK 879

Db 574 KPSAFEPALMOKSVNPKALELKNQOTLRADQMPFSESKQKVRENSWDSLSLRETVSQK 633

Qy 880 DVCVPKATHQEMDKISGKLESDTSLSKILDTVHSCERARELQKDHCEQRTGKMEQMKKK 939

Db 634 DVCVPKATHQEMDKISGKLESDTSLSKILDTVHSCERARELQKDHCEQRTGKMEQMKKK 693

Qy 940 FCVLKKLSEAKEIKSQLENQKVWEQELCSVRLTLNOEEKRRNADILNEKIREELGRI 999

Db 694 FCVLKKLSEAKEIKSQLENQKVWEQELCSVRLTLNOEEKRRNADILNEKIREELGRI 753

Qy 1000 EEOHREKLEVKQOOLQALRIODIELKSVESNLNQVSHTHENENYLLHNCMLKKEIAMLK 1059

Db 754 EEQHRKEVKKQLEQALRIQDIIEKSVESNLQVSHTHENENYLLHENCMLKKEIAMLK 813
Qy 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 1119
Db 814 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 873
Qy 1120 TMLTSKLKQKQDKEILEAEIESHPRLASAVODHDQIVTSRKSQEPAPHIAGDACLQRM 1179
Db 874 TMLTSKLKQKQDKEILEAEIESHPRLASAVODHDQIVTSRKSQEPAPHIAGDACLQRM 933
Qy 1180 NVDVSSIIYNNVHLQPLSEARQSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 1239
Db 934 NVDVSSIIYNNVHLQPLSEARQSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 993
Qy 1240 KEAEMHTQNEODNVNKHTEQOESLDQKLFQLOSKNMMLQOOLVHAHKKADNKSITIDIH 1299
Db 994 KEAEMHTQNEODNVNKHTEQOESLDQKLFQLOSKNMMLQOOLVHAHKKADNKSITIDIH 1053
Qy 1300 FLERKMOHLLKKEKNEEIFNYYNNHKNRIYQYEKEKAETENS 1341
Db 1054 FLERKMOHLLKKEKNEEIFNYYNNHKNRIYQYEKEKAETENS 1095

RESULT 15

ABUJ37747
ID ABUJ37747 standard; protein; 1095 AA.

AC ABUJ37747;

XX 15-MAY-2003 (first entry)

DT Human tumour-related protein - SEQ ID No 493.

DE Human: vaccine; gene therapy; T cell stimulation; T cell expansion;
KW tumour; breast cancer; cancer; immune response stimulation.

XX Homo sapiens.

XX WO200283956-A1.

XX 24-OCT-2002.

XX 15-APR-2002; 2002WO-US012378.

XX 13-APR-2001; 2001US-00834759.

XX 07-DEC-2001; 2001US-00007805.

XX 13-FEB-2002; 2002US-00076622.

XX (CORI-) CORIXA CORP.

XX Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;

XX Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;

XX Vedvick TS, McNeill PD, Durham M;

XX WPI; 2003-103376/09.

XX New polypeptide and polynucleotide useful for stimulating and/or

XX expanding T cells specific for a tumor protein and treating breast

XX cancer.

XX Disclosure; Page 310-312; 375pp; English.

XX The invention comprises a method of stimulating and/or expanding T cells

XX specific for a tumour protein. The invention further comprises human

XX nucleic acids and proteins that are associated with tumours (e.g. breast

XX cancer). The method and sequences of the invention are useful for

XX stimulating and/or expanding T cells specific for a tumour protein,

XX detecting the presence of cancer, stimulating an immune response in a

XX patient and treating breast cancer. The present amino acid sequence

XX represents a human tumour-related protein

XX Sequence 1095 AA;

Query Match 74.1%; Score 5130; DB 6; Length 1095;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 340 MSPAKETSEKTTAAKGRPRKIAWEKKEPTVKTCGVARVTSNKTQVLEKGRSKMTACPTK 399
Db 94 MSPAKETSEKTTAAKGRPRKIAWEKKEPTVKTCGVARVTSNKTQVLEKGRSKMTACPTK 153
Qy 400 ESSTKASANDORFPSEKQDEEYSCDSRSLFESSAKIQVCIPESIQKWEINREVEE 459
Db 154 ESSTKASANDORFPSEKQDEEYSCDSRSLFESSAKIQVCIPESIQKWEINREVEE 213
Qy 460 PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADPMPFPPEKQKDYEENSWSLCECTV 519
Db 214 PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADPMPFPPEKQKDYEENSWSLCECTV 273
Qy 520 SQKDVCLPKATHQKEIDKINGKLEESPNKDGILLKATCGMKVSIPTKALELMDQMTFKABP 579
Db 274 SQKDVCLPKATHQKEIDKINGKLEESPNKDGILLKATCGMKVSIPTKALELMDQMTFKABP 333
Qy 580 PGKPSAFEPATEMOKSVPNKALELKNQETLRADILPSEKQKDYEENSWSLCECTVS 639
Db 334 PGKPSAFEPATEMOKSVPNKALELKNQETLRADILPSEKQKDYEENSWSLCECTVS 393
Qy 640 QKDVCLPKAAHQKEIDKINGKLEESPVKDGILLKANCCKMKSIPTKALELMDQMTFKABPP 699
Db 394 QKDVCLPKAAHQKEIDKINGKLEESPVKDGILLKANCCKMKSIPTKALELMDQMTFKABPP 453
Qy 700 EKPSAFEPATEMOKSVPNKALELKNQETLRADILPSEKQKDYEENSWSLCECTVSQ 759
Db 454 EKPSAFEPATEMOKSVPNKALELKNQETLRADILPSEKQKDYEENSWSLCECTVSQ 513
Qy 760 KDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQMTFKABPE 819
Db 514 KDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQMTFKABPE 573
Qy 820 KPSAFEPATEMOKSVPNKALELKNQETLRADPMPFPSEKQKDYEENSWSLCECTVSQ 879
Db 574 KPSAFEPATEMOKSVPNKALELKNQETLRADPMPFPSEKQKDYEENSWSLCECTVSQ 633
Qy 880 DVCVPKATHQKEMDKISGKLEDSTSLKILDTVHSCEARELEQDHCQRTGKMBQMKKK 939
Db 634 DVCVPKATHQKEMDKISGKLEDSTSLKILDTVHSCEARELEQDHCQRTGKMBQMKKK 693
Qy 940 FCVLKKKLSAKEITKSOLEKQKWEQELCSVRITLNOEERKRRNADILNEKIREELGRI 999
Db 694 FCVLKKKLSAKEITKSOLEKQKWEQELCSVRITLNOEERKRRNADILNEKIREELGRI 753
Qy 1000 EEQHRKEVKKQLEQALRIQDIIEKSVESNLQVSHTHENENYLLHENCMLKKEIAMLK 1059
Db 754 EEQHRKEVKKQLEQALRIQDIIEKSVESNLQVSHTHENENYLLHENCMLKKEIAMLK 813
Qy 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 1119
Db 814 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 873
Qy 1120 TMLTSKLKQKQDKEILEAEIESHPRLASAVODHDQIVTSRKSQEPAPHIAGDACLQRM 1179
Db 874 TMLTSKLKQKQDKEILEAEIESHPRLASAVODHDQIVTSRKSQEPAPHIAGDACLQRM 933
Qy 1180 NVDVSSIIYNNVHLQPLSEARQSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 1239
Db 934 NVDVSSIIYNNVHLQPLSEARQSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 993
Qy 1240 KEAEMHTQNEODNVNKHTEQOESLDQKLFQLOSKNMMLQOOLVHAHKKADNKSITIDIH 1299
Db 994 KEAEMHTQNEODNVNKHTEQOESLDQKLFQLOSKNMMLQOOLVHAHKKADNKSITIDIH 1053
Qy 1300 FLERKMOHLLKKEKNEEIFNYYNNHKNRIYQYEKEKAETENS 1341
Db 1054 FLERKMOHLLKKEKNEEIFNYYNNHKNRIYQYEKEKAETENS 1095

Search completed: February 19, 2005, 02:04:43
Job time : 151.087 secs
